

基于注意力的多尺度水下图像增强网络

方 明^{②①} 刘小晗^{*①} 付飞蚬^①

^①(长春理工大学计算机科学技术学院 长春 130022)

^②(长春理工大学人工智能学院 长春 130022)

摘 要: 水下图像往往会因为光的吸收和散射而出现颜色退化与细节模糊的现象, 进而影响水下视觉任务。该文通过水下成像模型合成更接近水下图像的数据集, 以端到端的方式设计了一个基于注意力的多尺度水下图像增强网络。在该网络中引入像素和通道注意力机制, 并设计了一个多尺度特征提取模块, 在网络开始阶段提取不同层次的特征, 通过带跳跃连接的卷积层和注意力模块后得到输出结果。多个数据集上的实验结果表明, 该方法在处理合成水下图像和真实水下图像时都能有很好的效果, 与现有方法相比能更好地恢复图像颜色和纹理细节。

关键词: 水下图像增强; 深度学习; 注意力机制; 多尺度特征

中图分类号: TN911.73; TP391

文献标识码: A

文章编号: 1009-5896(2021)12-3513-09

DOI: 10.11999/JEIT200836

Multi-scale Underwater Image Enhancement Network Based on Attention Mechanism

FANG Ming^{②①} LIU Xiaohan^① FU Feiran^①

^①(School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China)

^②(School of Artificial Intelligence, Changchun University of Science and Technology, Changchun 130022, China)

Abstract: Due to the absorption and scattering, color degradation and detail blurring often occur in underwater images, which will affect the underwater visual tasks. A multi-scale underwater image enhancement network based on attention mechanism is designed in an end-to-end manner by synthesizing dataset closer to underwater images through underwater imaging model. In the network, pixel and channel attention mechanisms are introduced. A new multi-scale feature extraction module is designed to extract the features of different levels at the beginning of the network, and the output results are obtained via a convolution layer and an attention module with skip connections. Experimental results on multiple datasets show that the proposed method is effective in processing both synthetic and real underwater images. It can better recover the color and texture details of images compared with the existing methods.

Key words: Underwater image enhancement; Deep learning; Attention mechanism; Multi-scale features

1 引言

近年来, 各国均在积极发展海洋战略, 海洋信息的获取、传输和处理等关键技术对合理开发和利用海洋资源至关重要。水下图像是海洋信息的重要载体, 然而水下图像受复杂的水下环境和光照条件影响, 具有对比度低、纹理模糊、颜色失真、非均

匀光照和可视范围有限等问题(真实水下图像如图1所示)。对图像进行清晰化处理, 提高图像的质量, 有利于获取更多信息, 进一步进行水下目标跟踪和检测等任务。

目前在水下图像增强方面已经提出了很多传统的方法, 如直方图均衡化^[1]、颜色校正和对比度增强、基于融合的方法^[2]等。由于水下图像接近陆地的雾图, 因此基于去雾暗通道先验的方法^[3]提出了一些适用于水下的去雾和增强算法^[4-7]。常用的水下物理模型是由大气散射模型发展过来的, 考虑了不同波长的光在水下的衰减程度不同。后期, Akkaynak等人^[8]研究发现, 目前的水下成像模型忽略了一些关键的组成部分, 因而提出了修正的水下模型。然

收稿日期: 2020-09-27; 改回日期: 2021-04-25; 网络出版: 2021-07-14

*通信作者: 刘小晗 liuxh928@163.com

基金项目: 山东省支持青岛海洋科学与技术试点国家实验室重大科技专项(2018SDKJ0102-6)

Foundation Item: The Marine S&T Fund of Shandong Province for Pilot National Laboratory for Marine Science and Technology (Qingdao)(2018SDKJ0102-6)

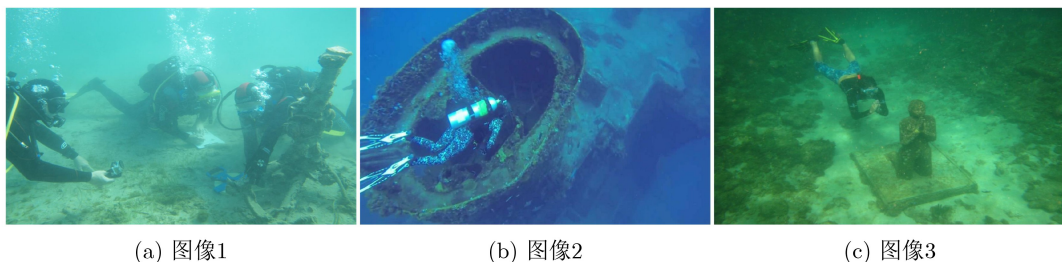


图1 水下图像

而,完善后的模型由于需要考虑各类参数的复杂性而限制了其广泛应用^[9]。

由于深度神经网络可以实现复杂非线性系统的端到端建模^[10],可用其拟合退化的水下图像到清晰水下图像之间的复杂映射关系,替代传统方法的复杂物理参数估计。因此,陆续有研究者提出深度学习方法在水下图像增强任务中的新方法。Hashisho等人^[11]使用CycleGAN生成成对的水下数据集,并使用UNet进行训练。但是使用CycleGAN训练时间较长,针对水体影响导致模糊较严重区域,颜色恢复效果不理想。Li等人^[12]提出了基于融合的卷积神经网络,制作的UIEBD数据集的图像分别经过白平衡、直方图均衡化、伽马校正后输入到卷积神经网络中进行特征融合。该方法能够很好地恢复出水下物体的颜色和细节,但是需要先对退化的水下图像进行处理,对于水体影响较大区域颜色恢复效果依然不理想。Fabbri等人^[13]使用生成对抗网络的方法,但是处理后的图像边缘相对于原图比较模糊。Islam等人^[14]同样基于生成对抗网络,提出了提升视觉感知的快速水下图像增强,达到了实时的效果。Zhang等人^[15]在生成对抗网络的基础上进行了改进,引入梯度差损失来锐化图像避免图像模糊,该方法提升了图像的清晰度及细节特征,但是恢复的水下图像有些颜色失真。Park等人^[16]提出了自适应加权的多鉴别器CycleGAN网络,在CycleGAN的基础上增加了内容鉴别器,保留输入图像的内容,并使用自适应加权方法限制两种鉴别器的损失,输出结果能够很好地恢复图像的颜色信息,但是在细节恢复上还可以进一步提升。

一般来说,基于深度学习的水下图像增强的重点在于数据集的获取,以及设计有效恢复水下图像细节和颜色的网络结构。针对这两个问题,本文提出了更接近真实水下场景的数据集合成方法,并提出了一个端到端的基于注意力的多尺度水下图像增强网络。对合成水下数据集和真实水下数据的主观与客观评价指标分析表明,本文提出的方法相比于其他水下图像增强方法,颜色和细节恢复得更好。本文的主要贡献:

(1)使用水下成像模型以及水体参数合成了一组新的数据集,该数据集同时考虑了水深和场景到相机的距离,在参数上更具有随机性,生成的水下图像具有多样性,更接近水下的复杂环境;

(2)在网络中引入像素和通道注意力机制,使网络更加关注水体影响较大区域的像素和更重要的通道信息;

(3)设计了一个多尺度特征提取模块,该模块使用不同采样率的平滑空洞卷积得到不同尺度的特征,再将这些特征通过注意力机制进行加权,以达到对不同尺度的特征采取不同的关注度的目的。

2 基于注意力的多尺度水下图像增强网络

2.1 数据集合成

目前有监督的网络在广泛的视觉任务中取得了显著的效果。然而,水下图像增强任务不同于一些高级视觉任务有公开的大型数据集。现有相关文献所使用的水下数据集大致分为3类:(1)使用NYU-v2 RGBD中的室内场景图及深度图按照水下成像模型进行合成^[17-19]; (2)使用CycleGAN^[11,20]的无监督方式用清晰图像生成符合水下风格的图像,从中挑选出效果好的图像对作为数据集,这种方法训练时间较长,需要人工去挑选效果好的图像,耗时耗力,并且很难控制模拟出各类水体特征;(3)UIEBD数据集^[12],由真实水下场景图像,通过12种现有方法进行处理,从中挑选出效果最好的图片作为对应的参考图。但是该数据集仅有890对图像,对于深度学习网络训练数量不足,很难训练出泛化能力强的网络。

因此,本文基于NYU-v2 RGBD数据集,利用水下成像模型重新合成水下数据集用于网络训练。相比于其他使用该数据集合成的方法,本文的数据集同时考虑了水深和场景到相机距离对图片的影响,更接近水下真实场景,同时采用随机性更高的水体参数,特征的覆盖能力更好,水下场景更具有多样性与复杂性。

由于水下场景与陆地雾天场景接近,为了描述水下物体的成像过程,常使用改进的雾霾成像公式

进行修正, 得到水下成像模型^[21], 如式(1)、式(2)所示

$$I_{\lambda}(x) = J_{\lambda}(x) \cdot T_{\lambda}(x) + (1 - T_{\lambda}(x)) \cdot B_{\lambda},$$

$$\lambda \in \{\text{red, green, blue}\} \quad (1)$$

$$T_{\lambda}(x) = \text{Nrer}(\lambda)^{d(x)} \quad (2)$$

其中, x 是水下场景的一点, $I_{\lambda}(x)$ 是相机拍摄到的水下图像, $J_{\lambda}(x)$ 是场景点的辐照度, $T_{\lambda}(x)$ 为中间能量比, 表示从水下场景中的点 x 反射后到达相机的亮度百分比, $d(x)$ 为场景到相机的距离, B_{λ} 是均匀的全局背景光, $\text{Nrer}(\lambda)$ 为归一化剩余能量比, 即每单位传播距离的剩余能量与初始能量之比。

由于光在水下会发生衰减, 因此光照通过深度为 $D(x)$ 的水体到达场景 x 经过衰减, 如式(3)所示

$$E_{\lambda}^{\text{obj}} = E_{\lambda}^{\text{in}} \cdot C_{\lambda} \quad (3)$$

$$C_{\lambda} = \text{Nrer}(\lambda)^{D(x)} \quad (4)$$

其中, E_{λ}^{obj} 表示由于初始能量 E_{λ}^{in} 的垂直吸收和散射而引起的物体的辐照度。 C_{λ} 代表波长衰减系数, 也是造成水下图像颜色偏色的主要原因。水下场景点 x 的辐照度 $J_{\lambda}(x)$ 等于环境光的照射量, 因此,

$$J_{\lambda}(x) = E_{\lambda}^{\text{obj}} \cdot \rho_{\lambda}(x) \quad (5)$$

其中, $\rho_{\lambda}(x)$ 表示波长为 λ 的光在场景点 x 的反射系数, 为了简化模型, 本文假设 $\rho_{\lambda}(x) = 1$ ^[22]。简化后的水下成像模型可以表示为

$$I_{\lambda}(x) = E_{\lambda}^{\text{in}} \cdot C_{\lambda} \cdot T_{\lambda}(x) + (1 - T_{\lambda}(x)) \cdot B_{\lambda},$$

$$\lambda \in \{\text{red, green, blue}\} \quad (6)$$

由于不同水体的衰减系数不同, 根据式(6)以及Zhou等人^[23]提供的参数合成不同水体(1), (2), (3)中成像的水下图像, 其中参数如表1所示。

该组参数根据真实水下图像对比选取了3组更接近真实情况的水体参数, 在参数中加入了随机数, 使数据更具有多样性。NYU-v2数据集提供了清晰图片 E_{λ}^{in} , 以及场景深度 $d(x)$ 。之后, 考虑到第(3)种水体衰减较大, 水深过深时图像几乎看不到物体, 因此在前两种随机设置 $D(x) \in [0.5, 10]$, 第(3)种 $D(x) \in [0.1, 2]$ 。最终合成了一个包含不同水体类型的图像数据集, 共7245对图像。部分水下图像视觉效果如图2所示, 图2(a)为原图, 图2(b)—图2(d)分别为水体类型(1), (2), (3)中的图像。从主观视觉效果上看本文生成的水下图像很接近真实水下图

表 1 不同水体中红(R), 绿(G), 蓝(B)通道中参数设置^[23]

水体类型	参数	R	G	B
(1)	$\text{Nrer}(\lambda)$	$0.79+0.06\text{rand}()$	$0.92+0.06\text{rand}()$	$0.94+0.05\text{rand}()$
	B_{λ}	$0.05+0.15\text{rand}()$	$0.60+0.30\text{rand}()$	$0.70+0.29\text{rand}()$
(2)	$\text{Nrer}(\lambda)$	$0.71+0.04\text{rand}()$	$0.82+0.06\text{rand}()$	$0.80+0.07\text{rand}()$
	B_{λ}	$0.05+0.15\text{rand}()$	$0.60+0.30\text{rand}()$	$0.70+0.29\text{rand}()$
(3)	$\text{Nrer}(\lambda)$	0.67	0.73	0.67
	B_{λ}	0.15	0.80	0.75



(a) 原图

(b) 水体类型 (1)

(c) 水体类型 (2)

(d) 水体类型 (3)

图 2 本文合成的水下图像

像,并且本文依据水下物理模型进行合成,符合水下成像原理。因此在真实水下数据集难以获得的情况下,使用该数据集进行深度学习训练是合理的。

2.2 网络结构

本文的目的是设计一个高效且鲁棒的端到端网络,该网络使用合成数据集进行训练,同时对于不同水体中得到的偏色模糊的真实水下图像也有着非常好的效果,在恢复图像颜色的同时保留了图像细节,使图像在整体视觉效果上更接近陆地拍摄清晰的图像。

网络结构如图3所示,由3部分组成:(1)基础网络;(2)多尺度特征提取模块;(3)注意力机制模块。其中,整体网络参数如表2所示。基础网络由两个CINR单元(卷积层+InstanceNorm层+ReLU层),2个下采样和CINR单元,残差块组(本文采用3个残差块),2个反卷积和CINR单元,1个CINR单元组成。其中在下采样和反卷积阶段,将相同大小的特征层进行连接。为了使网络能够对各种大小图片均有很好的效果,扩大特征提取的感受野,在基础网络的前面增加1个多尺度特征提取模块(Multi-scale Feature Extraction module, MFE),结构如图3(b)所示,该模块使网络能够关注到不同尺度的特征,从而增强整体的颜色与结构的恢复能力。最后引入了注意力机制模块(Attention mechanism, Attn),如图3(c)所示,该模块使网络能够对不同

通道特征、不同像素特征进行加权,关注到更重要的区域,如由水体导致的模糊区域。训练网络时使用平滑 L_1 损失和感知损失,使输出的图像与真实图像在颜色和内容上保持一致。

2.2.1 注意力机制模块(Attn)

为了更好地处理水下图像的颜色以及水体对图片区域的影响,本文同时使用通道注意力机制和像素注意力机制^[24],结构如图3(c)所示。该模块在处理不同类型的信息时提供了额外的灵活性,更加关注水下水体影响较大区域的像素和更重要通道信息。在通道注意力机制中,首先,利用全局平均池化将通道方向的全局空间信息转化为通道描述符。之后,特征通过1个全连接层,1个ReLU激活函数、1个全连接层、再加1个Sigmoid激活进行处理。

$$F'_c = \delta_s (\text{Conv} (\delta_r (\text{Conv} (\text{AvgPool} (F_c)))) \otimes F_c \quad (7)$$

其中, F_c 为前一阶段输出第 c 通道的特征, AvgPool 为平均池化操作, Conv 为卷积操作, δ_r , δ_s 分别为 ReLU 激活和 Sigmoid 激活函数。

接下来加入像素注意力机制是考虑到不同水下图像像素上的水体导致的模糊分布是不均匀的,使网络能够更加关注到水下由于光的衰减而看不清的区域,以及高频图像区域。

$$F_c^* = \delta_s (\text{Conv} (\delta_r (\text{Conv} (F'_c)))) \otimes F'_c \quad (8)$$

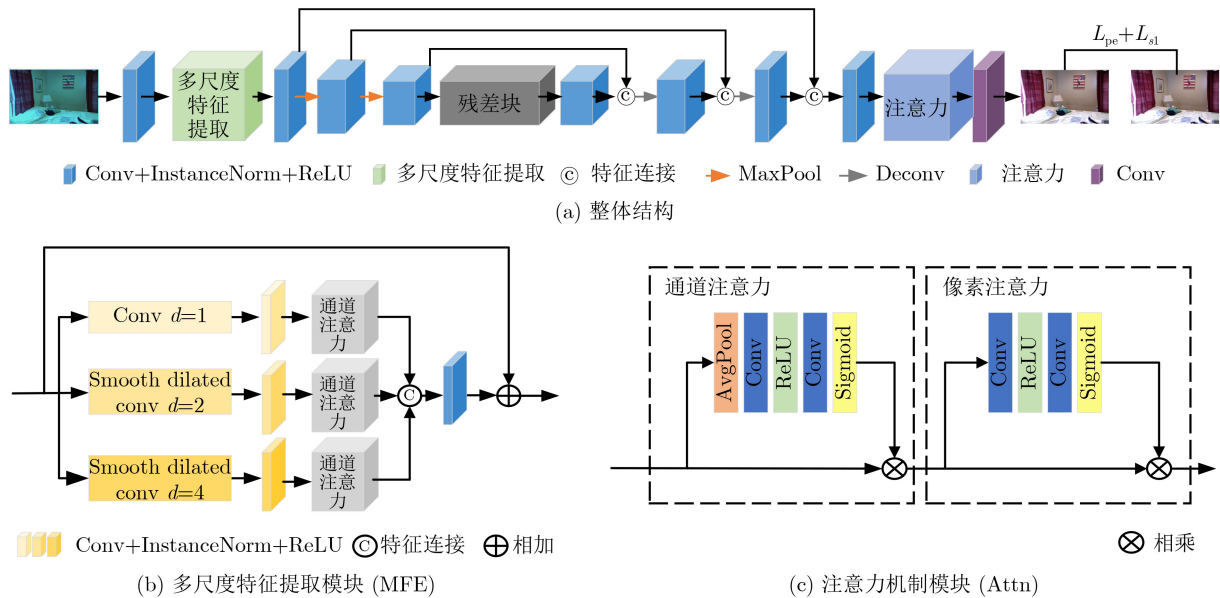


图3 网络结构图

表2 网络参数表

	CINR	MFE	CINR	Maxpool	CINR	Maxpool	CINR	ResBlocks	CINR	Deconv	CINR	Deconv	CINR	CINR	Attn	Conv
卷积核大小	3,1,1		3,1,1	2,2,0	3,1,1	2,2,0	3,1,1	3,1,1	3,1,1	4,2,1	3,1,1	4,2,1	3,1,1	3,1,1		3,1,1
输出通道数	64	64	64	64	128	128	256	256	256	128	128	64	64	64	64	64

2.2.2 多尺度特征提取模块(MFE)

在本文的网络中, 为了获取到不同层次的特征^[25], 设计了一个多尺度特征提取模块(图3(a)所示)。其中多尺度是指不同大小采样率 $rate = 1, 2, 4$ 的空洞卷积提取不同尺度范围的特征信息, 由于空洞卷积不同的扩张率, 感受野不同, 获得的特征区域信息大小不同。在1维情况下, 给定1-D输入 f , 空洞卷积的输出为

$$o(i) = \sum_{j=1}^k f[i + r \times j] \omega[j] \quad (9)$$

其中, r 为扩张率, k 为感受野大小, 当 $r = 1$ 时, 扩张的卷积对应于标准卷积。当 $r > 1$ 时, 扩张卷积即在标准卷积的卷积核中每两个相邻权重之间插入 $r - 1$ 个0, 感受野的大小由 k 扩张为 $(k - 1) \times r + 1$, 不同大小的感受野可以获取到不同范围的信息。考虑到原始的空洞卷积, 当采样率大于1时, 输出特征层中相邻的特征值与输入特征上是完全独立的, 会导致图像出现伪影。本文采用文献^[26,27]提出的平滑空洞卷积, 如图4所示。下一层的4个不同的点通过不同的颜色来指示, 可以看出, 它们是由上一层完全不相关的点获得的, 这将可能导致网格伪影。它在扩展卷积之前增加了一个额外的可分离共享卷积层, 从而增加了输入单元之间的依赖性。不同采样率的平滑空洞卷积得到的特征通过2.2.1节的注意力机制模块将不同尺度的特征进行权重分配, 再将特征连接到一起, 经过一层卷积, 自适应地提取不同尺度的特征, 得到的特征再与输入进行相加, 输向下一阶段。

2.2.3 损失函数

本文使用平滑的 L_1 损失和感知损失^[28]训练网络。其中, L_1 损失用来在训练过程中约束生成图像与参考图的像素值一致, 由于 L_1 范数可以防止潜在的梯度爆炸, 而平滑的 L_1 损失对于离群点更加鲁棒, 相比于 L_2 损失函数, 其对离群点、异常值的敏感度较低。平滑的 L_1 损失(L_{s1})可以表示为

$$L_{s1}(y', y) = \sum_i F_S(y'_i, y_i) \quad (10)$$

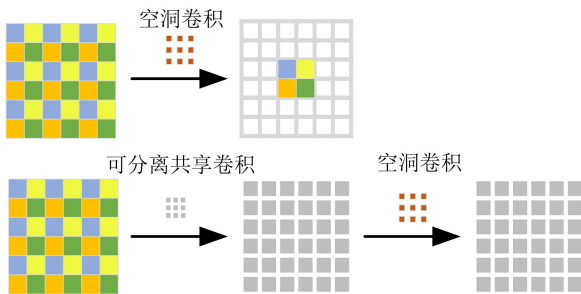


图4 平滑空洞卷积

其中

$$F_S(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{其他} \end{cases} \quad (11)$$

感知损失用来度量内容的感知, 利用从预先训练的深度神经网络来提取的多尺度特征, 将参考图像卷积得到的特征图与生成图像卷积得到的特征图作比较, 使得高层信息, 也就是内容和全局结构接近。本文使用预先训练的VGG16^[29]作为损失网络, 从前3个阶段的最后一层提取特征。感知损失(L_{pe})表示为

$$L_{pe}(y', y) = \sum_{j=1}^3 \frac{1}{C_j H_j W_j} \|\varphi_j(y') - \varphi_j(y)\|_2^2 \quad (12)$$

其中, $\varphi_j(y')$, $\varphi_j(y)$ ($j = 1, 2, 3$)为生成图像和真实图像在第 j 个阶段提取到的特征, C, W, H 为特征图的维度。总损失函数为平滑的 L_1 损失与感知损失的加权和, 如式(13)所示, λ 为感知损失的权重系数。

$$L = L_{s1} + \lambda L_{pe} \quad (13)$$

3 实验

3.1 训练设置

按照前文所提出的方法进行水下数据集生成。并随机选取6882张图片作为训练集。训练过程中, 随机截取 256×256 大小的区域作为输入。Batch size为16, 初始学习率设为0.01, 一共150个epoch, 每30个epoch, 学习率乘以0.5。感知损失权重系数 λ 为0.04。

本研究实验GPU环境为: Intel(R) Core(TM) i9-7960X CPU @ 3.30 GHz, Ubuntu18.04, 64位操作系统, 运行内存32 GB, NVIDIA Corporation GV100 [TITAN V], CUDA10.1, cuDNN 7.6.0, Python 3.6。CPU环境为: Intel(R) Xeon (R) CPU E3-1220 V2 @ 3.10 GHz, Windows10。

3.2 评价指标

本文在合成数据集上使用常用的PSNR和SSIM^[30]作为评价指标。在真实数据集上, 由于没有真实值对比, 因此选择使用水下图像质量测量UIQM^[31]作为评价指标。UIQM由3个水下测量指标组成: 水下图像色彩测量UICM, 水下图像清晰度测量UISM, 水下图像对比度测量UIconM。计算公式为

$$UIQM = c_1 \times UICM + c_2 \times UISM + c_3 \times UIconM \quad (14)$$

其中, 根据文献^[31]设置 $c_1 = 0.0282$, $c_2 = 0.2953$, $c_3 = 3.5753$ 。

3.3 合成水下图像评估

本文与其他方法的效果进行了比较, 对比的方

法包括水下暗通道先验(UDCP^[32]), 基于图像模糊和光吸收的水下图像恢复(IBLA^[33]), 基于快速场景深度估计模型的水下图像恢复(ULAP^[34]), 基于融合方法的水下图像增强(DUIENet^[12]), 基于GAN网络的快速水下图像增强(FUnIE-GAN^[14])。本文随机选取了145张图像作为测试集。实验结果如图5所示, 图5(a)为合成的水下图像, 图5(b)为原始清晰图像, 图5(c)为UDCP方法处理效果, 可以看出颜色对比度有些过度, 图5(d)、图5(e)分别为IBLA和ULAP方法, 这两种方法针对该合成图像效果不明显, 图5(f)、图5(g)分别为DUIENet和FUnIE-GAN方法, 能够恢复一些颜色和细节, 但是还有模糊感, 图5(h)为本文方法, 结果最清晰也最接近真实图像。另外对每种方法进行PSNR和SSIM值计算, 测试集上的均值如表3所示。可以看出, 无论哪一个指标, 本文方法数值结果都最优。从主观视觉效果和客观数据结果来看, 本文的方法是有效的。

3.4 真实水下图像评估

为了验证本文方法在真实水下图像上的有效性, 本文对78张真实水下图像进行了实验。图6所示为部分图像的效果图对比, 图6(a)为真实水下图像, 图6(b)—图6(g)分别是UDCP, IBLA, ULAP, DUIENet, FUnIE-GAN以及本文方法的处理效果图。本文的方法相比于其他方法整体色调更接近陆地清晰图像, 消除了水下图像由于水体原因造成的偏色情况, 效果对比度更高, 色彩更加鲜艳。在第1幅图中, 几乎完全去除了蓝色的水体遮挡, 恢复了水下物体的真实颜色, 同时细节也更加清晰。第2幅图中, 本文方法恢复的图像边缘更清晰, 对比度更高; 第3幅图中本文方法颜色对比度更接近陆地图像; 第4幅图中, 鱼的视觉感官更加清晰, 而且地面颜色恢复也更好。每张图像的UIQM数值结果如表4所示, 加粗为最优值。

本文对78张图片均进行了处理并计算了UIQM

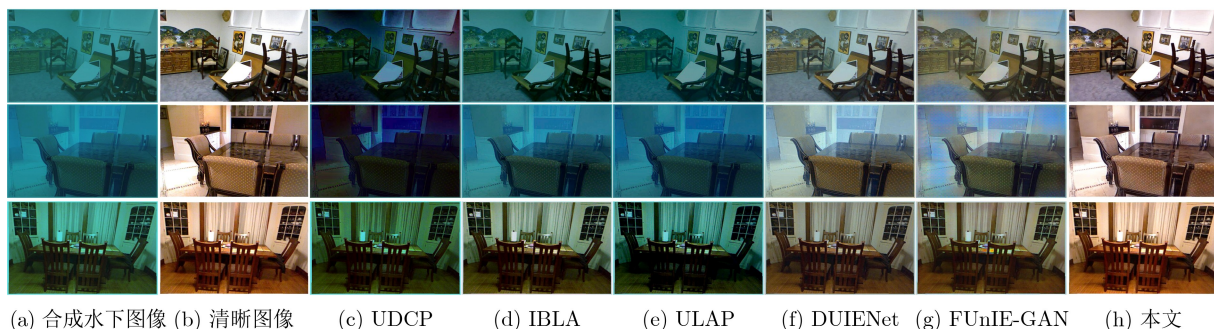


图5 不同方法在合成数据集上的处理效果

表3 不同方法在合成数据集上的PSNR和SSIM值

	UDCP	IBLA	ULAP	DUIENet	FUnIE-GAN	本文
PSNR	10.3134	14.8764	14.3421	15.6875	16.7233	28.4583
SSIM	0.5022	0.7225	0.7058	0.7889	0.7601	0.9110

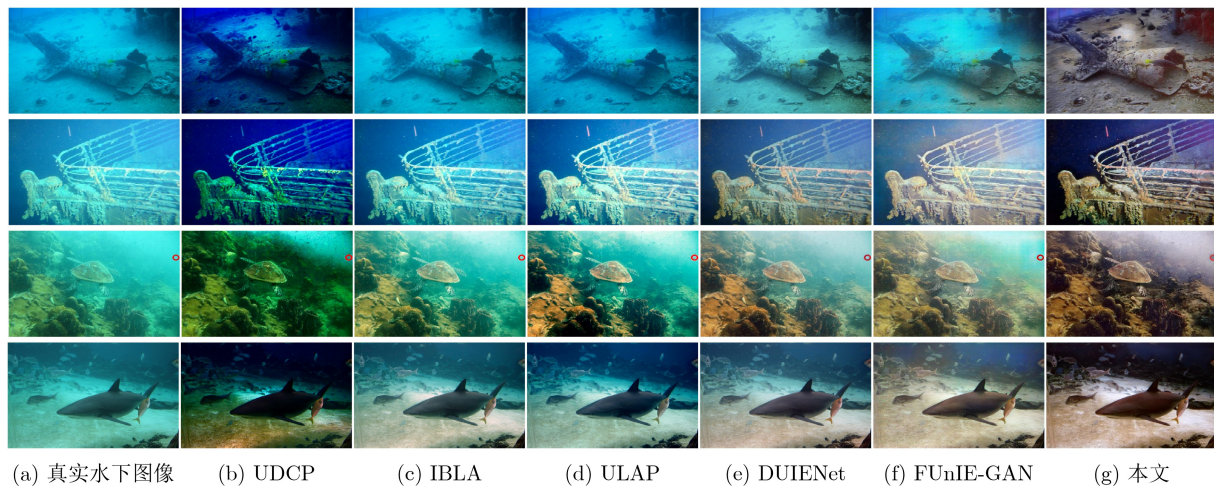


图6 不同方法在真实数据集上的处理效果

数值，其中UIQM的值为UICM, UISM, UIConM的加权和。对78张图片取平均值，每种方法的数值结果如表5所示。可以看出本文方法得到了最优结果。图7细节放大图的比较，图7(a)—图7(c)分别为DUIENet, FUnIE-GAN和本文方法。本文方法细节表现更加清晰。

3.5 消融实验

为了验证网络的有效性，本文在145张测试集

上进行了消融实验，验证注意力机制模块(Attn)、多尺度特征提取模块(MFE)以及感知损失的有效性。结果如表6所示。存在注意力机制模块、多尺度特征提取模块和感知损失情况下取得了最高的评优值。

3.6 运行时间

本文对每种方法的运行时间进行了统计，结果如表7所示。该结果为在CPU上处理1张大小为

表 4 图6中图片的不同方法的UIQM值

	图像	UDCP	IBLA	ULAP	DUIENet	FUnIE-GAN	本文
UIQM	第1幅	3.5135	3.8244	3.8666	4.1049	4.3301	4.8344
	第2幅	3.3235	3.5484	3.6567	3.7669	3.3144	3.9077
	第3幅	3.3387	3.2753	3.1166	3.7438	3.9245	4.2782
	第4幅	3.7411	4.1643	3.8395	4.3814	4.4317	4.6122

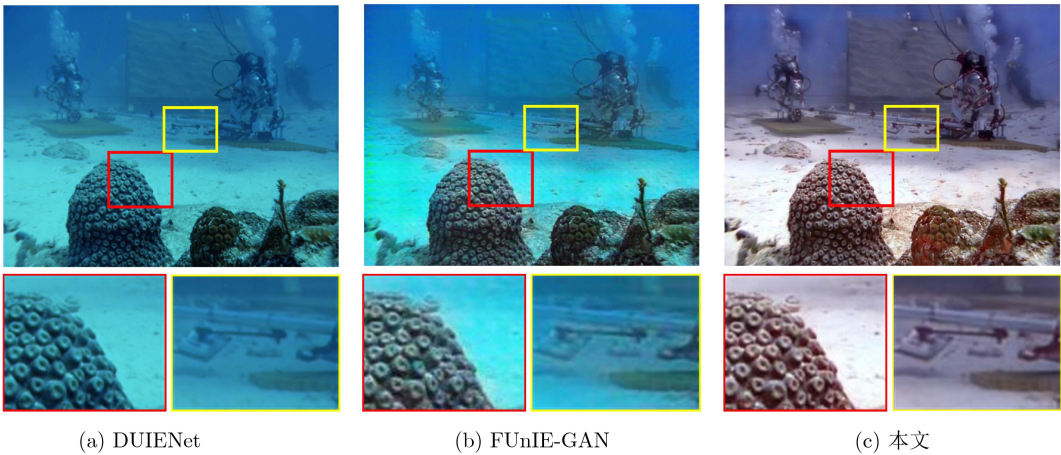


图 7 细节放大图

表 5 不同方法在真实数据集上的UIQM值

	UDCP	IBLA	ULAP	DUIENet	FUnIE-GAN	本文
UICM	3.8090	3.8621	3.8424	3.5210	4.2954	2.3348
UISM	4.4584	4.7460	4.8385	5.2352	5.2672	5.1937
UIConM	0.7023	0.5915	0.6202	0.6888	0.6779	0.7278
UIQM	3.9349	3.6251	3.7548	4.1079	4.1002	4.2016

表 6 网络结构消融实验数值结果

	MFE	Attn	感知损失	PSNR	SSIM
无MFE, Attn模块			✓	27.8182	0.9057
无Attn模块	✓		✓	27.8408	0.9092
无感知损失	✓	✓		28.1272	0.8967
本文完整网络	✓	✓	✓	28.4583	0.9110

表 7 不同方法运行时间表(s)

	UDCP	IBLA	ULAP	DUIENet	FUnIE-GAN	本文
运行时间	16.81	34.75	2.23	15.30	1.01	15.93

480 × 640 的图像所需要的时间。可以看出本文方法虽然不是最快的,但是结果与处理效果较好的深度学习方法DUIENet相当。

4 结束语

本文设计了一个基于注意力的多尺度水下图像增强网络,网络中引入了注意力机制,包括通道注意力和像素注意力,让网络能够更加关注到由于水体导致的模糊区域,从而更好地恢复水下图像,同时设计了一个多尺度特征提取模块,使网络能够更好地提取不同尺度的特征。使用水下图像成像模型以及水体参数合成了更加真实的数据集进行训练,训练好的网络不仅在合成数据集上获得了很好的效果,对于真实水下图像也能够很好地还原场景的颜色及细节。对比实验结果表明,本文方法要优于现有的很多方法,对水体造成的偏色能够有效地去除,增强后的图像很好地还原了物体原有特征。但同时,本文方法时间复杂度略高,很难满足实时处理的要求,这也是需要进一步改进的方向。

参 考 文 献

- [1] ZUIDERVELD K. Contrast Limited Adaptive Histogram Equalization[M]. Graphics Gems IV, Amsterdam: Elsevier, 1994: 474–485.
- [2] 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 (CVPR), Providence, USA, 2012: 81–88.
- [3] HE Kaiming, SUN Jian, and TANG Xiaoou. Single image haze removal using dark channel prior[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(12): 2341–2353. doi: 10.1109/TPAMI.2010.168.
- [4] LI Chongyi, GUO Jichang, CONG Runmin, *et al.* Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior[J]. *IEEE Transactions on Image Processing*, 2016, 25(12): 5664–5677. doi: 10.1109/TIP.2016.2612882.
- [5] 杨爱萍, 郑佳, 王建, 等. 基于颜色失真去除与暗通道先验的水下图像复原[J]. 电子与信息学报, 2015, 37(11): 2541–2547. doi: 10.11999/JEIT150483.
- [6] YANG Aiping, ZHENG Jia, WANG Jian, *et al.* Underwater image restoration based on color cast removal and dark channel prior[J]. *Journal of Electronics & Information Technology*, 2015, 37(11): 2541–2547. doi: 10.11999/JEIT150483.
- [7] 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. doi: 10.1109/MCG.2016.26.
- [7] REN Wenqi, LIU Si, ZHANG Hua, *et al.* Single image dehazing via multi-scale convolutional neural networks[C]. The 14th European Conference on Computer Vision, Amsterdam, The Kingdom of the Netherlands, 2016: 154–169.
- [8] AKKAYNAK D and TREIBITZ T. A revised underwater image formation model[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 6723–6732.
- [9] ANWAR S and LI Chongyi. Diving deeper into underwater image enhancement: A survey[EB/OL]. <https://arxiv.org/abs/1907.07863>, 2019.
- [10] 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. doi: 10.1109/LRA.2017.2730363.
- [11] HASHISHO Y, ALBADAWI M, KRAUSE T, *et al.* Underwater color restoration using u-net denoising autoencoder[EB/OL]. <https://arxiv.org/abs/1905.09000>. 2019.
- [12] LI Chongyi, GUO Chunle, REN Wenqi, *et al.* An underwater image enhancement benchmark dataset and beyond[EB/OL]. <https://arxiv.org/abs/1901.05495>, 2019.
- [13] FABBRI C, ISLAM M J, and SATTAR J. Enhancing underwater imagery using generative adversarial networks[C]. 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 2018: 7159–7165.
- [14] ISLAM M J, XIA Y Y, and SATTAR J. Fast underwater image enhancement for improved visual perception[J]. *IEEE Robotics and Automation Letters*, 2020, 5(2): 3227–3234. doi: 10.1109/LRA.2020.2974710.
- [15] ZHANG Tingting, LI Yujie, and TAKAHASHI S. Underwater image enhancement using improved generative adversarial network[J]. *Concurrency and Computation Practice and Experience*, 2020: e5841. doi: 10.1002/cpe.5841.
- [16] PARK J, HAN D K, and KO H. Adaptive weighted multi-discriminator CycleGAN for underwater image enhancement[J]. *Journal of Marine Science and Engineering*, 2019, 7(7): 200. doi: 10.3390/jmse7070200.
- [17] ANWAR S, LI Chongyi, and PORIKLI F. Deep underwater image enhancement[EB/OL]. <https://arxiv.org/abs/1807.03528>, 2018.
- [18] LI Chongyi, ANWAR S, and PORIKLI F. Underwater scene prior inspired deep underwater image and video enhancement[J]. *Pattern Recognition*, 2020, 98: 107038. doi: 10.1016/j.patcog.2019.107038.
- [19] SILBERMAN N, HOIEM D, KOHLI P, *et al.* Indoor

- segmentation and support inference from RGBD images[C]. The 12th European Conference on Computer Vision, Florence, Italy, 2012: 746–760.
- [20] 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, Italy, 2017: 2242–2251.
- [21] CHIANG J Y and CHEN Y C. Underwater image enhancement by wavelength compensation and dehazing[J]. *IEEE Transactions on Image Processing*, 2012, 21(4): 1756–1769. doi: [10.1109/TIP.2011.2179666](https://doi.org/10.1109/TIP.2011.2179666).
- [22] DING Xueyan, WANG Yafei, YAN Yang, *et al.* Jointly adversarial network to wavelength compensation and dehazing of underwater images[EB/OL]. <https://arxiv.org/abs/1907.05595>, 2019.
- [23] ZHOU Yuan and YAN Kangming. Domain adaptive adversarial learning based on physics model feedback for underwater image enhancement[EB/OL]. <https://arxiv.org/abs/2002.09315>, 2020.
- [24] QIN Xu, WANG Zhilin, BAI Yuanchao, *et al.* FFA-Net: Feature fusion attention network for single image dehazing[EB/OL]. <https://arxiv.org/abs/1911.07559>, 2019.
- [25] 柳长源, 王琪, 毕晓君, 等. 基于多通道多尺度卷积神经网络的单幅图像去雨方法[J]. *电子与信息学报*, 2020, 42(9): 2285–2292. doi: [10.11999/JEIT190755](https://doi.org/10.11999/JEIT190755).
- LIU Changyuan, WANG Qi, BI Xiaojun, *et al.* Research on rain removal method for single image based on multi-channel and multi-scale CNN[J]. *Journal of Electronics & Information Technology*, 2020, 42(9): 2285–2292. doi: [10.11999/JEIT190755](https://doi.org/10.11999/JEIT190755).
- [26] WANG Zhengyang and JI Shuiwang. Smoothed dilated convolutions for improved dense prediction[C]. The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 2018: 2486–2495.
- [27] CHEN Dongdong, HE Mingming, FAN Qingnan, *et al.* Gated context aggregation network for image dehazing and deraining[C]. 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, USA, 2019: 1375–1383.
- [28] JOHNSON J, ALAHI A, and LI Feifei. Perceptual losses for real-time style transfer and super-resolution[C]. 14th European Conference on Computer Vision, Amsterdam, The Kingdom of the Netherlands, 2016: 694–711.
- [29] SIMONYAN K and ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]. International Conference on Learning Representations (ICLR), San Diego, USA, 2015: 1–5.
- [30] 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. doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861).
- [31] PANETTA K, GAO C, and AGAIAN S. Human-visual-system-inspired underwater image quality measures[J]. *IEEE Journal of Oceanic Engineering*, 2016, 41(3): 541–551. doi: [10.1109/JOE.2015.2469915](https://doi.org/10.1109/JOE.2015.2469915).
- [32] DREWS P JR, DO NASCIMENTO E, MORAES F, *et al.* Transmission estimation in underwater single images[C]. 2013 IEEE International Conference on Computer Vision Workshops, Sydney, Australia, 2013: 825–830.
- [33] PENG Y T and COSMAN P C. Underwater image restoration based on image blurriness and light absorption[J]. *IEEE Transactions on Image Processing*, 2017, 26(4): 1579–1594. doi: [10.1109/TIP.2017.2663846](https://doi.org/10.1109/TIP.2017.2663846).
- [34] SONG Wei, WANG Yan, HUANG Dongmei, *et al.* A rapid scene depth estimation model based on underwater light attenuation prior for underwater image restoration[C]. The 19th Pacific-Rim Conference on Multimedia, Hefei, China, 2018: 678–688.
- 方 明: 男, 1977年生, 副教授, 博士, 硕士生导师, 研究方向为图像处理、计算机视觉。
- 刘小晗: 女, 1995年生, 硕士生, 研究方向为图像处理、计算机视觉。
- 付飞舫: 男, 1989年生, 博士, 研究方向为图像处理、计算机视觉。

责任编辑: 马秀强