

基于多阶段协同处理的水下图像增强

袁红春, 赵华龙*, 高凯

上海海洋大学信息学院, 上海 201306

摘要 设计了一种多阶段水下图像增强模型, 可以同时将空间精细纹理和高级上下文信息两种特征融合。模型由三个阶段组成, 前两个阶段采用编码器-解码器结构, 第三阶段则采用并行注意力子网, 所提模型可以同时学习空间细节和上下文信息两种特征, 并且引入了监督注意力模块, 能够加强特征学习, 还设计了一个跨阶段特征融合机制用来巩固前后子网的中间特征。最后将所提模型与其他水下增强模型在同一测试集上运行, 从运行结果得出, 所提模型处理后的水下图像在主观视觉效果和客观评价质量上均优于大部分对比算法, 在 Test-1 测试集上, 峰值信噪比和结构相似度分别达到了 26.2962 dB 和 0.8267。

关键词 图像处理; 水下图像增强; 多阶段; 空间细节; 监督注意力

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP231422

Underwater Image Enhancement Based on Multi-Stage Collaborative Processing

Yuan Hongchun, Zhao Hualong*, Gao Kai

College of Information Technology, Shanghai Ocean University, Shanghai 201306, China

Abstract We propose a multi-stage underwater image enhancement model that can simultaneously fuse spatial details and contextual information. The model is structured in three stages: the first two stages utilize encoder-decoder configurations, and the third entails a parallel attention subnet. This design enables the model to concurrently learn spatial nuances and contextual data. A supervised attention module is incorporated for enhanced feature learning. Furthermore, a cross-stage feature fusion mechanism is designed is used to consolidate the intermediate features from preceding and succeeding subnets. Comparative tests with other underwater enhancement models demonstrate that the proposed model outperforms most extant algorithms in subjective visual quality and objective evaluation metrics. Specifically, on the Test-1 dataset, the proposed model realizes a peak signal-to-noise ratio of 26.2962 dB and structural similarity index of 0.8267.

Key words image processing; underwater image enhancement; multi-stage; spatial details; supervise attention

1 引言

近年来,随着计算机技术的快速发展,计算机视觉系统被广泛应用于很多水下场景,如海洋环境监测、渔业养殖、水下考古等。水下图像承载了重要的海洋信息,但受到水下浮游粒子和光照等因素影响,水下图像会产生清晰度下降、颜色偏差、亮度偏暗等问题。而这些扭曲和模糊的图像会成为计算机执行视觉任务的障碍,所以增强水下图像、获得高质量的水下信息,成为开展水下研究任务的先决条件。

早前,很多研究人员采用非物理模型的方法对图像进行处理,比如传统的直方图均衡化^[1]、限制对比度

自适应直方图均衡化(Contrast limited adaptive histogram equalization, CLAHE)^[2]、融合方法^[3]和基于视网膜大脑皮层(Retinex)理论^[4]的增强方法等。这些方法没有考虑引起图像退化的各种因素,只是对图像本身的像素值进行调整以实现图像质量的增强。后来,Hitam等^[5]以 CLAHE 为基础设计了一种混合 CLAHE 算法,弥补了传统 CLAHE 方法的缺点,在对水下图像进行增强时不会产生额外的噪声。Naik等^[6]提出一种浅层神经网络模型,相比于其他模型,该模型不仅拥有更少的参数量,同时性能也十分优良,具有良好的泛化性。Zhou等^[7]以一种新的思路构建了自己的水下图像数据集,其中主要包含湖泊、池塘等浑浊水域

场景的水下图像数据,同时还提出了一种无监督的深度学习框架,可专注于提升浑浊水域中水下图像的质量,其定性和定量评价结果表明,该方法能有效地提高图像的清晰度。考虑到在同一水下场景中拍摄的图像通常具有相似的退化特性,Qi等^[8]提出了一种基于暹罗编解码器架构的水下图像协同增强网络,可以为图像之间的增强提供丰富的互补信息。Zamir等^[9]提出了一个用于图像渐进恢复的多阶段网络,通过融合编码器-解码器(Encoder-decoder)子网和原始分辨率两种模块来同时获得空间细节和高级上下文信息。Saleh等^[10]提出了一种新的无监督水下图像增强框架,主要模块分别是多色空间拉伸和概率自适应实例归一化,前者用来调整图像的颜色平衡,而后者则是通过捕获图像特征和编码来增强过程中的不确定性,结果表明该框架在性能上拥有优越的竞争力。Yang等^[11]提出了一种轻量级的自适应特征融合网络,该网络通过减少卷积通道来减少参数量,而且比其他先进模型的参数量更少、运行速度更快。Li等^[12]首次尝试在生成式对抗网络水下任务中混合两种输入,设计了一种新的融合对抗网络,该网络可以有效地校正颜色,产生视觉上令人愉悦的增强结果。Li等^[13]提出一种通过介质传输与引导且有多色空间嵌入的水下图像增强网络,将不同颜色空间的特征纳入到一个统一的结构中,从而丰富了特征表示的多样性,有效地提高水下图像的视觉质量。Liu等^[14]提出了一种基于监督学习的自适应学习注意网络(Adaptive learning attention network, LANet),可以保留浅层信息,并且能够自适应地学习水下图像的重要特征信息,在不同的水下数据集上都具有良好的性能。

Li等^[15]构建了水下图像增强卷积神经网络 WaterNet,通过实验对比分析,该模型对水下图像的增强效果优于其他大部分模型,此外,他们建立了新的水下图像数据集,为后续的水下图像增强研究提供了更多的数据参考。2021年,Wang等^[16]提出了一种使用RGB和LAB双色空间的水下图像增强深度学习框架,解决了在RGB颜色空间中图像偏暗和饱和度不高的问题。2022年,Jiang等^[17]提出了一种基于拉普拉斯金字塔的级联网络,在每个金字塔级的剩余网络上实现级联,以从粗到细的方式和显著降低的复杂度进行渐进预测,通过实验得出,该模型性能优异且能够对水下图像进行大幅增强。

分析这些现有的方法,发现很少有人尝试将多级设计用于水下图像增强,并找到其中阻碍性能提升的架构瓶颈。现有的多级设计大多采用编码器-解码器体系结构,该体系结构在编码宽泛的上下文信息方面行之有效,但在处理图像过程中却不能很好地保留空间图像细节。除此之外,一些设计采用单尺度的级联结构,这种结构可以提供准确的空间信息,但在语义输出上较差。所以,想要在多阶段架构的水下图像增强

中得到优质的输出结果,需要将两种设计结合起来。于是本文提出了一种用于水下图像增强的多阶段渐进恢复架构(Underwater multi-stage progressive image restoration, UWMPR)。主要贡献为

- 1) 一种全新的适用于水下图像增强的多阶段方法,将两种不同的子网络融合在一起进行特征互补,能够生成上下文信息丰富且空间信息准确的输出。
- 2) 提出了一种利用连接集成像素注意和通道注意的并行注意模块(Parallel attention module, PAM),主要关注照明特性和更重要的颜色信息。
- 3) 一个有效的监督注意模块(Supervised attention module, SAM),在每个阶段充分利用恢复的图像,在进一步传播之前细化传入的特征。
- 4) 一种跨阶段聚合多尺度特性的策略。

2 所提方法

2.1 网络体系结构

图1表示为UWMPR模型整体结构,其中, P_1 、 P_2 、 P_3 分别表示不同阶段的PAM, S_1 、 S_2 、 S_3 分别表示不同阶段的SAM。所提模型由3个阶段组成,每个阶段的第一步都是相同的,首先经过一个卷积层然后再经过一个PAM;在PAM之后,前两个阶段采用编码器-解码器结构对输入图像进行进一步增强,编码器-解码器子网拥有宽泛的感受域,能够对上下文信息进行有效的学习;最后一个阶段则采用并行注意子网(Parallel attention sub network, PASNet)进行操作,能够在最终的输出图像中保留所需的空间细节。

对于所提模型的3个阶段,并不是进行简单的级联操作,而是在每两个阶段之间加入一个SAM,用于增强参考图像的监督效果。此外,引入跨阶段特征融合(Cross-stage feature fusion, CSFF)模块,将前一阶段的特征传输到后一个阶段,以提升图像增强的效果。在输入图像上采用多图像块输入,将图像分割成4个图像块,并在模型的每个阶段输入图像块的特征信息:第一阶段采用4个图像块,第二阶段采用2个图像块,第三阶段采用原图像。

UWMPR模型的增强过程从最底层的阶段1开始,原始图像 B_3 被分割成4个不重叠的图像块 $B_{3,j}$, $j=1, \dots, 4$,随后将其输入模型,首先经过一个卷积层和 P_1 ,生成以下卷积特征

$$C_{3,j} = P_1[\text{Cnov}(B_{3,j})], j \in \{1, \dots, 4\}, \quad (1)$$

式中: $\text{Conv}(\cdot)$ 为卷积操作。之后再将 $C_{3,j}$ 输入编码器 η_3 生成以下特征

$$E_{3,j} = \eta_3(C_{3,j}), j \in \{1, \dots, 4\}. \quad (2)$$

然后沿着空间维度连接 $E_{3,j}$ 上相邻的特征得到 $E_{3,j}^*$

$$E_{3,j}^* = E_{3,2j-1} \oplus E_{3,2j}, j \in \{1, 2\}, \quad (3)$$

式中: \oplus 表示连接操作符。接下来,连接起来的特征 $E_{3,j}^*$ 经过解码器 ξ_3 得到 $D_{3,j}$, $D_{3,j}$ 经过 S_3 再加上阶段2

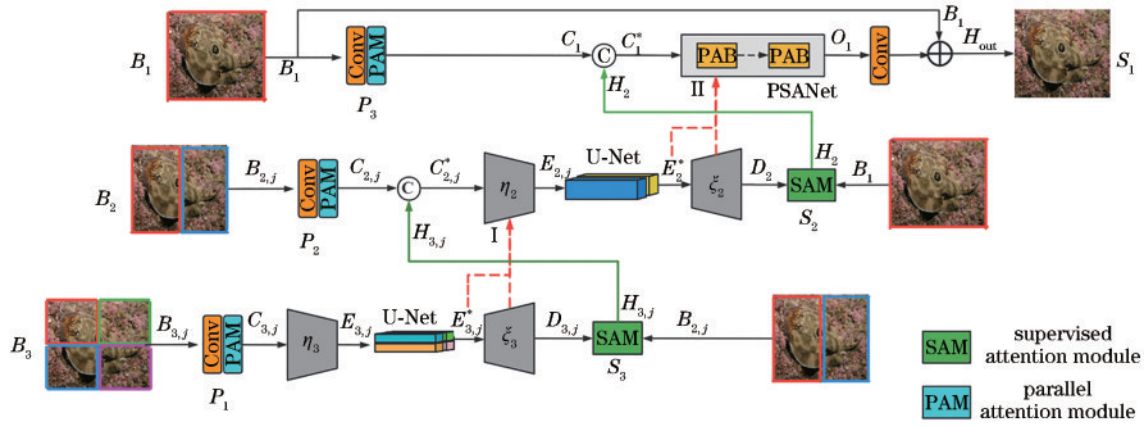


图1 UWMPR模型整体架构

Fig. 1 Overall architecture of UWMPR model

的图像块 $B_{2,j}, j=1, 2$ 作为输入, 得到 $H_{3,j}$,

$$D_{3,j} = \xi_3(E_{3,j}^*), j \in \{1, 2\}, \quad (4)$$

$$H_{3,j} = S_3(D_{3,j}) + B_{2,j}, j \in \{1, 2\}. \quad (5)$$

在阶段 2 中, 图像块 $B_{2,j}$ 经过一个卷积层和 P_2 得到特征 $C_{2,j}$ 后, 首先要与阶段 1 中 SAM 的输出 $H_{3,j}$ 连接起来,

$$C_{2,j} = P_2[\text{Conv}(B_{2,j})], j \in \{1, 2\}, \quad (6)$$

$$C_{2,j}^* = C_{2,j} \oplus H_{3,j}, j \in \{1, 2\}. \quad (7)$$

在阶段 2 的编码器中, 采用跨阶段融合 I, 将特征 $C_{2,j}^*$ 与来自阶段 1 的编码器特征 $E_{3,j}^*$ 和解码器特征 $D_{3,j}$ 相融合, 一同作为阶段 2 编码器 η_2 的输入, 经过处理后得到特征 $E_{2,j}$

$$E_{2,j} = \eta_2[\text{Conv}(E_{3,j}^*) + \text{Conv}(D_{3,j}) + C_{2,j}^*], j \in \{1, 2\}, \quad (8)$$

然后沿着空间维度连接 $E_{2,j}$ 上相邻的特征得到 E_2^*

$$E_2^* = E_{2,1} \oplus E_{2,2}, \quad (9)$$

接下来, 连接起来的特征 E_2^* 经过解码器 ξ_2 得到 D_2, D_2 经过 S_2 再加上阶段 3 的图像块 B_1 作为输入, 得到 H_2 ,

$$D_2 = \xi_2(E_2^*), \quad (10)$$

$$H_2 = S_2(D_2) + B_1. \quad (11)$$

之后, 图像就进入模型最后的阶段 3, 在阶段 3 中, 图像块 B_1 经过一个卷积层和 P_3 得到特征 C_1 后, 首先要与阶段 2 中 SAM 的输出 H_2 连接起来得到总的特征 C_1^* ,

$$C_1 = P_3[\text{Conv}(B_1)], \quad (12)$$

$$C_1^* = C_1 \oplus H_2. \quad (13)$$

随后则进入第 3 阶段的子网络, 在这个阶段子网络不再是编码器-解码器, 而是换成了 PASNet 模块, 采用跨阶段融合 II, 将特征 C_1^* 与来自阶段 2 的编码器特征 E_2^* 和解码器特征 D_2 相融合, 一同作为 PASNet 模块的输入, 经过处理后得到特征 O_1 ,

$$O_1 = P_{\text{pas}}[\text{Conv}(E_2^*) + \text{Conv}(D_2) + C_1^*], \quad (14)$$

特征 O_1 在经历最后一个卷积层后, 与图像块 B_1 相加,

得到最后的输出 H_{out} ,

$$H_{\text{out}} = \text{Conv}(O_1) + B_1. \quad (15)$$

2.2 互补特征处理

编码器-解码器网络^[18]首先逐步将输入映射到低分辨率, 然后逐步应用反向映射来恢复原始分辨率。虽然这些模型有效地编码了多尺度信息, 但多次降采样会对空间细节产生损耗。PASNet 作为一个单尺度的串联结构, 虽然在提取空间细节上表现良好, 但由于接受域有限, 它的语义输出不那么完善, 这表明了上述架构设计存在选择上的固有局限性, 它们能够生成空间信息准确或上下文信息丰富的可靠输出, 但两者不能兼得。利用这两种设计的优点进行特征互补, 设计了一个多阶段的深度学习网络, 其中在前两个阶段包含了编码器-解码器子网, 最后阶段使用了一个 PASNet。

所提编码器-解码器子网络如图 2 所示, 它从传统的 U-Net^[19] 改进而来, 在模块中的降采样、上采样以及跳跃连接当中都加入了通道注意块, 对比普通的卷积层能够更好地提取图像中的特征, 从而增强模型的效果。

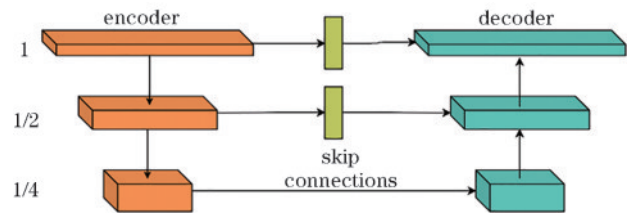


图2 U-Net编码器解码器子网

Fig. 2 Subnet of U-Net encoder and decoder

为了保留从输入图像到输出图像的细节, 所提网络在最后一阶段引入了 PASNet, 它不使用任何降采样操作, 并能生成空间丰富的高分辨率特征。它由多个并行注意块 (Parallel attention block, PAB) 组成, 每个块包含多个 PAM。PAB 的原理图如图 3 所示。



图 3 PASNet子网中的PAB
Fig. 3 PAB in PASNet subnet

每一个PAM都是由一个像素注意(Pixel attention, PA)块和一个通道注意(Channel attention, CA)块 并联组成,其具体构成方式如图4所示,其中IN (InstanceNorm2d)为归一化操作。

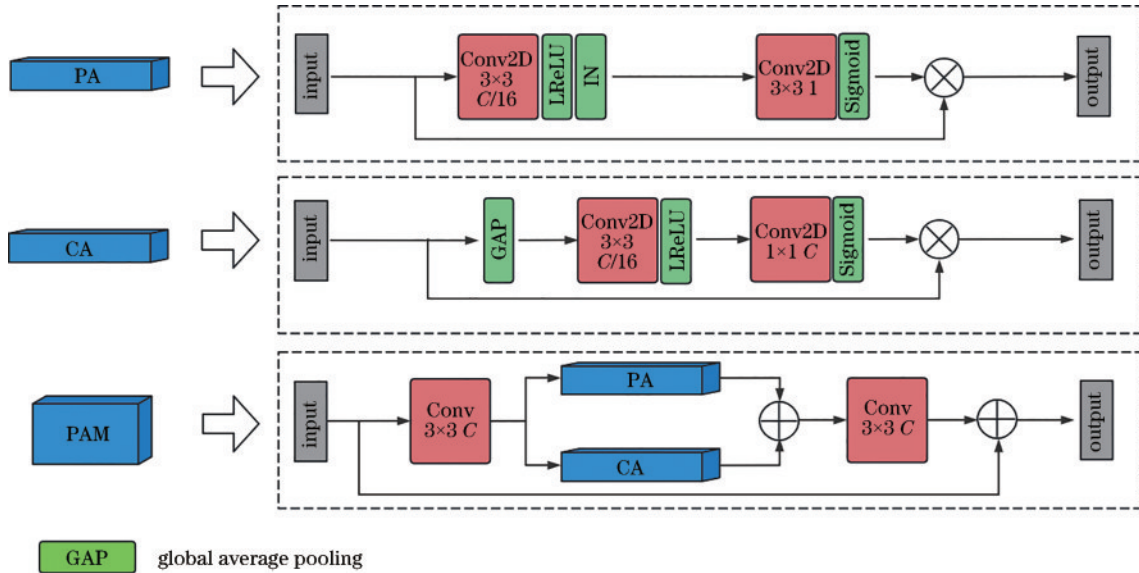


图 4 并行注意模块

Fig. 4 Parallel attention module

1) 像素注意块:设计的PA主要关注像素值。在特定范围内的像素值越高,表明图像具有越高的照度和越明亮的颜色。首先,使用两个卷积层来关注像素特征,特征的形状从 $C \times H \times W$ 变化到 $1 \times H \times W$,其中 $H \times W$ 为空间维, H 为图像的高, W 为图像的宽, C 为通道的数量。最后,将像素特征乘以 $PA \times PA$ 的输入特征。PA输出特性的通道数为 C ,定义为64个。

2) 通道注意块:CA主要为每个通道分配一个权重,平衡颜色信息,调整不同通道间不均匀的照明特征。首先,使用全局平均池化将特征纳入信道转换,特征的形状从 $C \times H \times W$ 变化为 $C \times 1 \times 1$ 。然后,使用两个卷积层来学习权重值。最后,将权重乘以CA的输入特征。

许多水下图像增强方法^[20-21]都忽略了照明特征和颜色信息之间的相关性,所以所提方法设计了PAM来重点关注照明特征和颜色信息。首先,使用一个降采样来提取特征。然后,利用并行的PA和CA来学习相应的特征,将学习到的特征以像素为单位添加至网络中。此外,还增加了一个卷积层来平衡照明特征和颜色信息。最后,引入了主要关注实际信息的局部残差连接。

2.3 CSFF

跨阶段融合机制如图5所示。在模型的框架中,

引入了CSFF模块,在模型的第一阶段到第二阶段,即两个编码器-解码器子网之间,将第一阶段的编码器和解码器提取的特征分别经过一个 1×1 卷积细化,然后传播到第二阶段,与第二阶段的编码器特征进行聚合。在第二阶段到第三阶段,即编码器-解码器子网与PASNet子网之间,同样是先将第二阶段的特征经过一个 1×1 卷积进行细化,然后与PASNet子网特征进行聚合。所提出的CSFF模块包含以下几个优点: 1)在编码器-解码器中重复使用上采样和降采样操作,使网络信息不容易遭受损失;2)将前一个阶段的特征传输到后一阶段,可以丰富后一个阶段提取的特征;3)能使图像增强系统的整体框架更加稳定。

2.4 SAM

最近用于水下图像增强的多阶段网络^[22-23]在每两个阶段之间都是直接传递。但是,所提网络在模型的每两个阶段之间引入一个SAM,有助于提升模型的整体效果。SAM的原理如图6所示,其贡献有两个方面:1)它为每个阶段的图像增强都提供了有效的参考图像;2)在SAM中会生成注意力图,会自适应地筛选重要的特征信息,并且抑制包含信息不多的特征。

如图6所示,在输入 F_{in} 进入SAM的早期阶段,首先经过一个 1×1 的卷积层生成一个残差图像 R_s , S 为阶段数。将残差图像添加到退化的输入图像 I 中,得

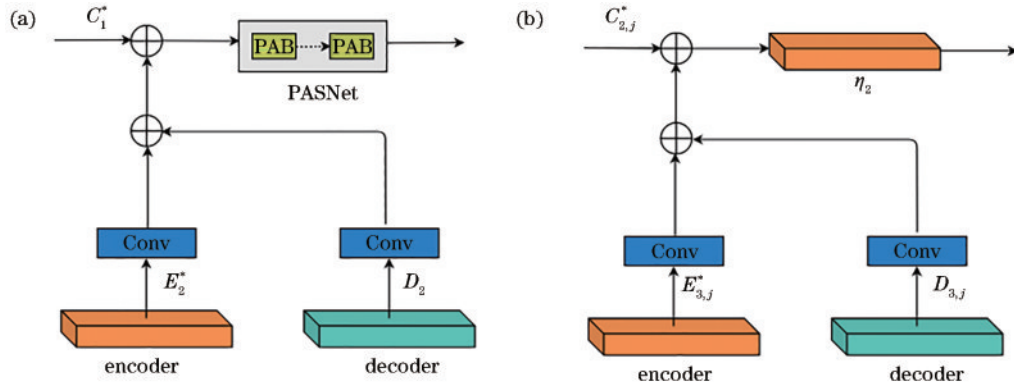


图 5 跨阶段融合机制。(a)跨阶段融合 I ;(b)跨阶段融合 II

Fig. 5 Cross-stage fusion mechanism. (a) Cross-stage fusion I ; (b) cross-stage fusion II

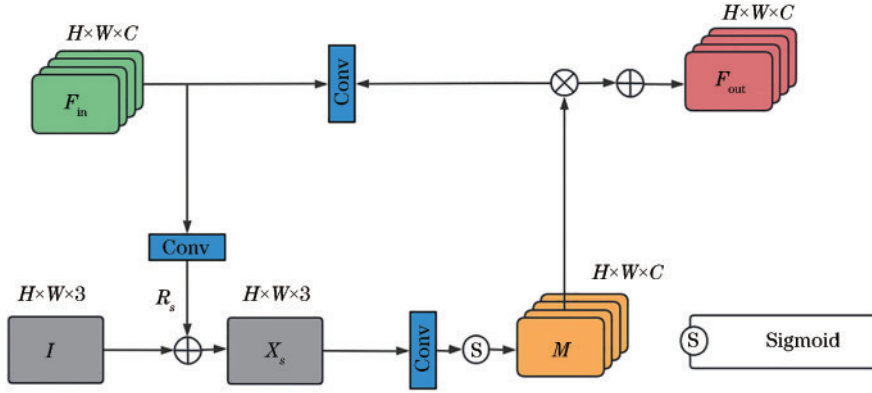


图 6 SAM原理示意图

Fig. 6 Principle diagram of SAM

到增强后的图像 X_s 。增强的图像 X_s 接下来再通过一个 1×1 卷积层和激活函数, 得到注意力地图 M 。与此同时, 在上面的路径中, 输入图像 F_{in} 也经过了一个 1×1 卷积层, 而在下面的路径中得到的注意力地图 M 正好回归, 与增强过后的 F_{in} 融合。最后, 将 SAM 产生的注意增强特征表示 F_{out} 传递到下一阶段进行进一步处理。

2.5 损失函数

为了实现良好的视觉效果, 使用 Charbonnier (char) 损失和边缘 (edge) 损失来优化模型的训练。在任意的阶段 S , 模型没有直接得到增强图像 X_s , 而是经过一个 1×1 的卷积得到图像 R_s , 添加原始图像 I 得到: $X_s = I + R_s$ 。损失函数的表达式为

$$L = \sum_{S=1}^3 [L_{char}(X_s, Y) + \lambda L_{edge}], \quad (16)$$

式中: Y 为参考图像; λ 为用来控制两个损失函数的权重, 在实验中设置为 0.05。char 损失函数与一般的均方误差相比, 添加了一个常量 ϵ , 使得函数在零点处也可导, 在实验中 ϵ 设置为 10^{-3} 。char 损失函数的表达式为

$$L_{char} = \sqrt{\|X_s - Y\|^2 + \epsilon^2}, \quad (17)$$

式中: $\|\cdot\|$ 表示取范数。

边缘损失则是考虑结构和纹理信息, 因为这些信息与人的主观感知效果密切相关, 可以带来更好的视觉效果。其表达式为

$$L_{edge} = \frac{\sum_{i=1}^w \sum_{j=1}^h E_{0i,j} \cdot |Y_{i,j} - X_{i,j}|}{W_1 H_1}, \quad (18)$$

式中: W_1 和 H_1 分别为图像的宽和高; $|Y - X|$ 为参考图像与输入图像的绝对误差; E_0 为对应的参考图像 Y 的边缘映射, 通过 MATLAB 的 Canny 边缘检测器对 Y 进行训练得到。Canny 基于双阈值方法, 能够更好地实现边缘选取, 提取出真正的图像边缘。边缘损失采用 E_0 加权后, 边缘部分的占比就会变大, 那么这个损失函数就能更好地优化边缘信息。

3 分析与讨论

3.1 实验设置

所提模型在 PyTorch 的深度学习框架中实现, 计算机的 GPU 为 NVIDIA 的 GeForce RTX3090 (24 GB)。采用的 UWMPR 是端到端可训练的, 不需要预先培训。所提模型在编码器-解码器的每个尺度上使用 2 个 CA 块为了降低采样次数, 采用了 2×2 的最大池化操作并设置步幅为 2。在最后一个阶段, 采用了含有 3 个 PAB 的 PASNet, 而每个 PAB 又进一步采用了

8个PAM。训练时,批处理大小为4,迭代次数为500。所提模型采用Adam优化器^[24],将初始学习速率设置为 5×10^{-5} ,使用余弦退火策略^[25],将它稳步减少至 1×10^{-6} 。

3.2 数据集

构建了训练数据集UIMP,其中包含了2000张成对的水下图像,包括真实水下图像和参考图像。UIMP首先从Underwater imagenet^[26]中提取800对真实水下图像及参考图像,随后从UIEBD、EUVP^[27]、UFO-120^[28]三个数据集中分别提取400对、300对、300对共计2000张水下图像。测试数据集包括两部分,Test-1和Test-2。Test-1是从Underwater imagenet数据集(去除训练集所挑选的800对图像)再次挑选的60张水下退化图像;除此之外,为了更好地测试模型性能,采用水下摄像机在学校池塘中拍摄了一组真实水下图像数据集,从中挑选出了39张图片作为第二部分Test-2,因为没有对应的参考图像所以将Test-2用作挑战集。

3.3 对比试验

3.3.1 定量评价的结果

为了比较所有方法的性能,对每个数据集都进行了定量数据评价和主观视觉质量评价。在测试集Test-1中采用结构相似性(Structural similarity, SSIM)与峰值信噪比(Peak signal-to-noise ratio, PSNR)来评价增强后图像与参考图像之间的差距。同时引入水下图像质量度量(Underwater image quality measure, UIQM)与清晰度测量指标(Underwater image sharpness measure, UISM)评价图片本身的质量,这4种指标得分越高说明图像增强的效果越好。表1为选取的8种模型LANet, Water-Net, UDCP (Underwater dark channel prior), CLAHE, UCM (Unsupervised colour correction method), RGHS (Relative global histogram stretching), HE (Histogram equalization), UDnet (Uncertainty distribution network), ICM (Integrated colour model)与所提模型在Test-1上的实验数据对比。用黑体和斜体表示的值分别表示最好的结果和第二好的结果。由表1可知,所提出的UWMPR在大多数情况下优于其他方法,只是在UISM上逊色于CLAHE方法。所提方法中使用的两个损失函数都偏向于空间细节和纹理结构,所以可以看到SSIM和PSNR两个评价指标都比较高,其指标与第二好的结果相比,相差也较大。但在UIQM和UISM两个指标上,所提方法与结果较好的方法差距不大,但相比于一般方法,得分较为可观。

对挑战集Test-2采用的评价指标中,除了UIQM之外,还添加了水下图像色彩度测量(Underwater image colorfulness measure, UICM),水下彩色图像质量评估(Underwater color image quality evaluation, UCIQE),自然图像质量(Natural image quality evaluator,

表1 不同方法在Test-1上的定量对比

Table 1 Quantitative comparison of different methods on Test-1

Method	SSIM ↑	PSNR ↑	UIQM ↑	UISM ↑
LANet	0.7618	24.0140	2.9936	7.1028
Water-Net	0.7738	22.1162	2.8657	6.1129
UDCP	0.5151	14.0129	1.8562	6.3351
CLAHE	0.6654	19.1925	2.9670	7.8442
UCM	0.7365	21.3801	2.8264	6.7269
RGHS	0.6908	22.1530	2.2372	5.8900
HE	0.5820	16.0704	2.9176	7.2327
ICM	0.7468	24.7373	2.6430	6.6378
UWMPR	0.8267	26.2962	3.1153	7.3001

NIQE)。其中UICM和UCIQE用于评价图像的色彩是否鲜艳,得分越高表示图像质量越好,而NIQE的得分越低表示图像的质量越好。从表2可以看出,所提出的UWMPR在大多数情况下优于其他方法,仅在UCIQE指标上排名第三。

表2 不同方法在Test-2上的定量对比

Table 2 Quantitative comparison of different methods on Test-2

Method	UIQM ↑	UICM ↑	UCIQE ↑	NIQE ↓
LANet	1.6654	1.4114	0.4990	3.1466
Water-Net	1.9609	2.3730	0.5744	3.5235
UDCP	1.2528	3.3168	0.5796	2.2941
CLAHE	2.6714	4.7339	0.5257	2.4418
UCM	2.5903	4.9081	0.6129	2.4354
RGHS	2.7410	5.2200	0.6328	2.4440
ICM	2.4386	3.7073	0.5249	2.3629
UDnet	2.3023	3.3325	0.5525	2.4193
UWMPR	2.8086	5.2840	0.5936	2.2816

3.3.2 定性评价的结果

图7为各种模型在测试集Test-1上运行得到的增强图像。对图像进行主观分析,可以得到以下结果:1)经过CLAHE和HE两个模型增强后的图像多数出现曝光现象,虽然亮度很高,但整体结构已经被破坏,在表1中,二者的SSIM和PSNR数据也在所有方法中处于较低水平;2)UCM模型整体泛红,但增强后的图像亮度很高;3)经过ICM、RGHS和UCDP方法处理后的图像偏绿,不过其中RGHS增强的图像在亮度上有一定提升,有更好的视觉效果,经过ICM和UDCP增强的图像则亮度一般;4)经过LANet增强后的图像整体效果不错,但图像上仿佛蒙了一层雾,这是水下图像常见的雾化现象,LANet并没有有效地解决该现象;5)Water-Net模型解决了上述现象,增强后的图像清晰度更高,但图像整体偏暗,亮度和对比度不够;6)相较于其他方法,经过所提方法增强后的图像不仅在色彩上更接近参考图像,而且对比度更高,视觉效果更好,整体清晰度更佳,与参考图像质量最为接近。

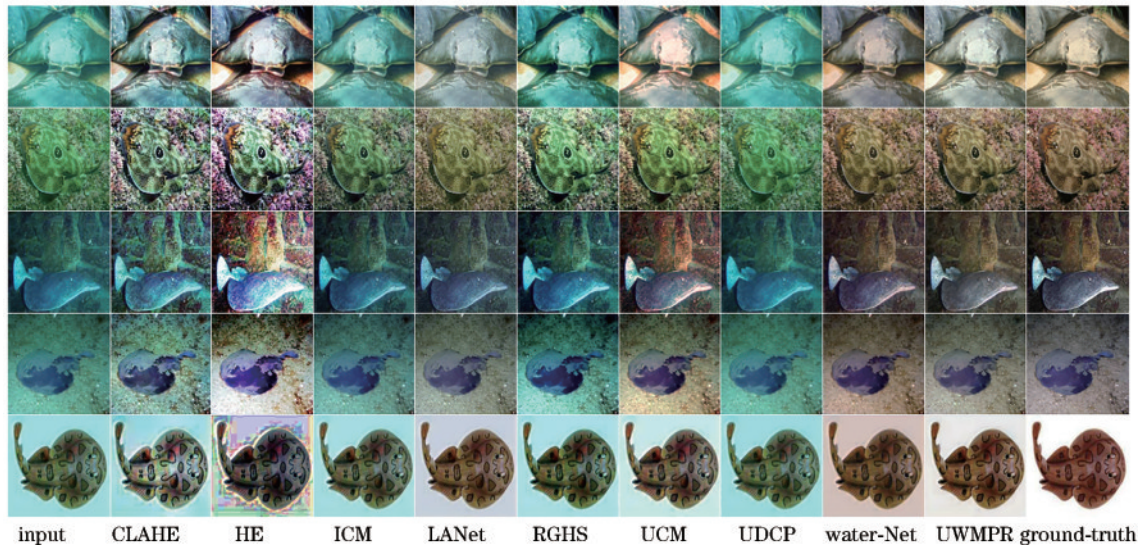


图 7 各种模型在 Test-1 上的视觉质量对比

Fig. 7 Comparison of visual quality of various models on Test-1

图 8 为各种方法在挑战集 Test-2 上得到的结果, 各种方法的优劣表现得更为直观。由图 8 可知: 1) 经过 LANet 和 Water-Net 增强的图像整体偏模糊, 其中 Water-Net 得到的图像还显示出伪影, 图像中增加了并不存在的鱼; 2) UDCP 效果整体偏暗, 如同在夜晚中一般; 3) CLAHE 效果较好, 清晰度也十分可观, 但整体颜色与原图像存在偏差; 4) UCM 增强的图像整体视觉效果欠佳, 但清晰度尚可; 5) RGHS、ICM 和 UDNet 的效果都较好, 其中 RGHS 增强后的图像对于雾化现象处理较好, ICM 和 UDNet 则清晰度一般, 有些模糊; 6) 所提方法不但解决了水环境下存在的雾化现象, 整体清晰度优异, 而且图像色彩均衡且与原图像的整体色彩最为契合, 亮度和对比度相比于其他方法更高, 在一些细节上处理得更加优秀。

3.4 消融实验

为了验证 UWMPR 模型中提出的 PAM、SAM 和 CSFF 模块的具体效果, 设计了针对性的消融实验进行性能分析, 并在 Test-1 和 Test-2 上进行定量分析。因为模型迭代次数越多, 耗费时间越长, 所以消融实验中模型的迭代次数均设置为 40。

3.4.1 PAM

对 PAM 的消融实验, 实际上是将模型中所有的 PAM 替换为普通的 CA 块。在同等条件下, 训练两个模型, 在两个数据集上的定量分析结果如表 3 和表 4 所示。PAM 设计初衷就是为了重点关注图像的亮度特征和颜色信息, 可以看到拥有 PAM 的模型在涉及到颜色、亮度、清晰度等指标时表现得更加出色, 数据有大幅提升。但作为代价, 含 PAM 的模型在 SSIM 和 UIQM 两个指标上略低于普通的 CA 块。

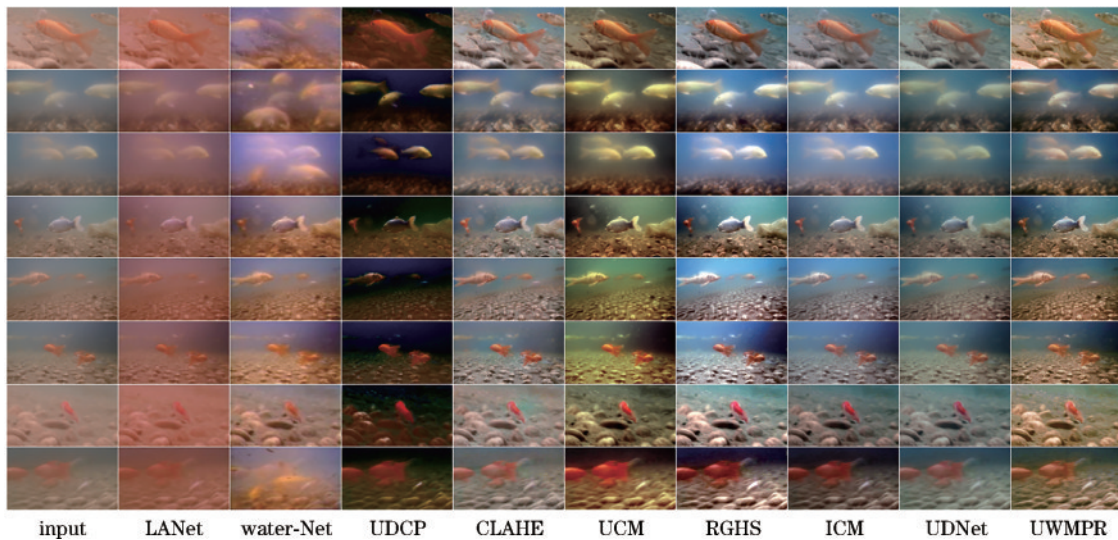


图 8 各种模型在 Test-2 上的视觉质量对比

Fig. 8 Comparison of visual quality of various models on Test-2

表 3 Test-1 上 PAM 消融实验数据

Table 3 Experimental data of PAM ablation on Test-1

Module	SSIM ↑	PSNR ↑	UIQM ↑	UICM ↑	UCIQE ↑	NIQE ↓
PAM	0.8113	25.8524	3.1630	7.4841	0.5671	5.2264
CA	0.8144	25.6719	3.1653	7.4389	0.5631	5.5143

表 4 Test-2 上 PAM 消融实验数据

Table 4 Experimental data of PAM ablation on Test-2

Module	UISM ↑	UCIQE ↑	NIQE ↓
PAM	5.7477	0.5936	2.2816
CA	4.8855	0.5760	2.6842

3.4.2 SAM 和 CSFF 模块

将 SAM 和 CSFF 模块的消融实验合并在一起,训练后得到的数据如表 5 和表 6 所示,其中最优和次优结果被标注为黑体和斜体。可以看到,除了 SSIM 和 NIQE 两个指标外,在其余度量指标上都是同时含有 SAM 和 CSFF 模块的模型效果最佳。在大多数指标上,同时不含有 SAM 和 CSFF 模块的模型效果次优或最优,由此可知单独含 SAM 或 CSFF 模块的并非最佳选择,各有缺陷,只有两者合并在一起才能形成互补,得到最佳效果。表 6 中 UCIQE 指标中同时含有 SAM 和 CSFF 模块的模型效果不佳,可能因为 Test-2 测试集是在池塘中采集,水域较浅,图像自身饱和度和对比度质量较高,色偏不严重,所以对 UCIQE 指标不太敏感,导致效果不佳。

表 5 Test-1 上 SAM 和 CSFF 模块消融实验数据

Table 5 Ablation experimental data of SAM and CSFF module on Test-1

Module	SSIM ↑	PSNR ↑	UIQM ↑	UICM ↑	UCIQE ↑	NIQE ↓
SAM√, CSFF√	0.8113	25.8524	3.1630	7.4841	0.5671	5.2264
SAM×, CSFF√	0.8142	25.7833	3.1115	7.3723	0.5653	5.2483
SAM√, CSFF×	0.8122	25.8135	3.1461	7.4524	0.5638	5.1009
SAM×, CSFF×	0.8142	25.3971	3.1586	7.4293	0.5663	5.0551

表 6 Test-2 上 SAM 和 CSFF 模块消融实验数据

Table 6 Ablation experimental data of SAM and CSFF module on Test-2

Module	UISM ↑	UCIQE ↑	NIQE ↓
SAM√, CSFF√	6.1157	0.5936	2.2816
SAM×, CSFF√	5.8129	0.5920	2.4856
SAM√, CSFF×	5.7477	0.5986	2.4541
SAM×, CSFF×	5.8619	0.5937	2.3609

通过以上消融实验可知,UWMPR 模型中提出的 PAM、SAM 和 CSFF 模块在提升增强效果的同时,也降低了 SSIM 指标。这是因为这些模块为了提升效果,不断地向图像中融合提取出的特征,一定程度上破坏了图像自身的结构,造成了结构相似性指标的降低。

4 结 论

在这项工作中,提出了一个用于水下图像增强的多阶段的图像增强模型。通过融合编码器-解码器子网和 PASNet,使模型能够同时对图像的空间细节信息和高级上下文信息进行学习;通过在每两个阶段之间加入一个 SAM,提高对原始图像特征的保留程度;通过引入跨阶段的特征融合机制,巩固前后子网的上下文中间特征,对模型的性能作了进一步的提升。通过定量和定性的实验分析,证明了经过 UWMPR 增强后的图像,整体清晰度优异,图像色彩更加均衡且与原图像的整体色彩更加契合,亮度和对比度更高,在一些细节上处理得更加优秀,并且在一定程度上解决了水下环境存在雾化现象这一问题。采用多阶段架构方式,模型规格较大,因此如何降低模型量级而不使增强效果下降是接下来考虑展开的研究。

参 考 文 献

- [1] Zuiderveld K. Contrast limited adaptive histogram equalization[M]//Heckbert P S. Graphics gems. Amsterdam: Elsevier, 1994: 474-485.
- [2] Reza A M. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement[J]. Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology, 2004, 38(1): 35-44.
- [3] Zou W, Wang X, Li K Q, et al. Self-tuning underwater image fusion method based on dark channel prior[C]//2016 IEEE International Conference on Robotics and Biomimetics (ROBIO), December 3-7, 2016, Qingdao, China. New York: IEEE Press, 2017: 788-793.
- [4] AbuNaser A, Doush I A, Mansour N, et al. Underwater image enhancement using particle swarm optimization[J]. Journal of Intelligent Systems, 2015, 24(1): 99-115.
- [5] Hitam M S, Awalludin E A, Yussof W N J H W, et al. Mixture contrast limited adaptive histogram equalization for underwater image enhancement[C]//2013 International Conference on Computer Applications Technology (ICCAT), January 20-22, 2013, Sousse, Tunisia. New York: IEEE Press, 2013.
- [6] Naik A, Swarnakar A, Mittal K. Shallow-UWnet: compressed model for underwater image enhancement (student abstract)[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(18): 15853-15854.
- [7] Zhou W H, Zhu D M, Shi M, et al. Deep images enhancement for turbid underwater images based on unsupervised learning[J]. Computers and Electronics in Agriculture, 2022, 202: 107372.

- [8] Qi Q, Zhang Y C, Tian F, et al. Underwater image co-enhancement with correlation feature matching and joint learning[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, 32(3): 1133-1147.
- [9] Zamir S W, Arora A, Khan S, et al. Multi-stage progressive image restoration[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 14816-14826.
- [10] Saleh A, Sheaves M, Jerry D, et al. Adaptive uncertainty distribution in deep learning for unsupervised underwater image enhancement[EB/OL]. (2022-12-18) [2023-02-03]. <https://arxiv.org/abs/2212.08983>.
- [11] Yang H H, Huang K C, Chen W T. LAFFNet: a lightweight adaptive feature fusion network for underwater image enhancement[C]//2021 IEEE International Conference on Robotics and Automation (ICRA), May 30-June 5, 2021, Xi'an, China. New York: IEEE Press, 2021: 685-692.
- [12] Li H Y, Zhuang P X. DewaterNet: a fusion adversarial real underwater image enhancement network[J]. *Signal Processing: Image Communication*, 2021, 95: 116248.
- [13] Li C Y, Anwar S, Hou J H, et al. Underwater image enhancement via medium transmission-guided multi-color space embedding[J]. *IEEE Transactions on Image Processing*, 2021, 30: 4985-5000.
- [14] Liu S B, Fan H J, Lin S, et al. Adaptive learning attention network for underwater image enhancement[J]. *IEEE Robotics and Automation Letters*, 2022, 7(2): 5326-5333.
- [15] 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*, 2019, 29: 4376-4389.
- [16] Wang Y D, Guo J C, Gao H, et al. UIEC²-Net: CNN-based underwater image enhancement using two color space[J]. *Signal Processing: Image Communication*, 2021, 96: 116250.
- [17] Jiang N F, Chen W L, Lin Y T, et al. Underwater image enhancement with lightweight cascaded network [J]. *IEEE Transactions on Multimedia*, 2022, 24: 4301-4313.
- [18] Thomas R, Thampi L, Kamal S, et al. Dehazing underwater images using encoder decoder based generic model-agnostic convolutional neural network[C]//2021 International Symposium on Ocean Technology (SYMPOL), December 9-11, 2021, Kochi, India. New York: IEEE Press, 2022.
- [19] 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.
- [20] Hong L, Wang X, Xiao Z L, et al. WSUIE: weakly supervised underwater image enhancement for improved visual perception[J]. *IEEE Robotics and Automation Letters*, 2021, 6(4): 8237-8244.
- [21] Sun X, Liu L P, Li Q, et al. Deep pixel-to-pixel network for underwater image enhancement and restoration[J]. *IET Image Processing*, 2019, 13(3): 469-474.
- [22] Huo F S, Li B H, Zhu X G. Efficient wavelet boost learning-based multi-stage progressive refinement network for underwater image enhancement[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), October 11-17, 2021, Montreal, BC, Canada. New York: IEEE Press, 2021: 1944-1952.
- [23] Liu J N, Liu Z L, Wei Y H, et al. Recovery for underwater image degradation with multi-stage progressive enhancement[J]. *Optics Express*, 2022, 30(7): 11704-11725.
- [24] Kingma D P, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2014-12-22) [2023-02-03]. <https://arxiv.org/abs/1412.6980>.
- [25] Loshchilov I, Hutter F. SGDR: stochastic gradient descent with warm restarts[EB/OL]. (2016-08-13) [2023-02-04]. <https://arxiv.org/abs/1608.03983>.
- [26] 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.
- [27] 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.
- [28] Islam M J, Luo P G, Sattar J. Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception[C]//Robotics: Science and Systems XVI, July 12-16, 2020, Corvallis, Oregon, USA. Robotics: Science and Systems Foundation, 2020.
- [29] Seif G, Androutsos D. Edge-based loss function for single image super-resolution[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 15-20, 2018, Calgary, AB, Canada. New York: IEEE Press, 2018: 1468-1472.