

DOI: 10.13382/j.jemi.B2306359

多特征选择与双向残差融合的无监督水下图像增强*

胡雨航 赵磊 李恒 刘辉

(昆明理工大学信息工程与自动化学院 昆明 650500)

摘要:如今,利用合成的成对数据集训练的有监督模型泛化能力弱,在多变的实际水下环境中表现不佳,而无监督模型虽摆脱了成对数据集的依赖,但生成图像可能因缺少特征信息导致图像视觉质量较差。故以循环生成对抗网络为架构,提出多特征选择与双向残差融合的水下图像增强方法。一方面,设计以混合注意力为基础的多特征选择模块对水下图像的多种特征进行选择,再由双向残差融合对传统U型跳跃连接进行优化,使图像特征高效表达,有效恢复水下图像的纹理与色彩。另一方面,在判别器中引入混合注意力并提出内容感知损失和风格感知损失,保证增强图像在全局内容、局部纹理、风格特征等方面和清晰图像一致。与现有的无监督和有监督模型相比较,该模型PSNR分别提高了6%和2%,SSIM分别提高了4%和3%,对水下图像有着显著的增强效果,在色彩真实度和饱和度上相比其他现有方法更加优秀。

关键词:无监督模型;循环生成对抗网络;多特征选择;双向残差融合;水下图像增强

中图分类号: TP391; TN919.8 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Unsupervised underwater image enhancement with multi-feature selection and bidirectional residual fusion

Hu Yuhang Zhao Lei Li Heng Liu Hui

(Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: Currently, the supervised models trained on synthetic paired datasets have weak generalization ability and perform poorly in diverse real underwater environments. Although unsupervised models are not dependent on paired datasets, the lack of feature information may result in the generated images with poor visual quality. Therefore, with the architecture of cyclic generation adversarial networks, the underwater image enhancement method of multi-feature selection and bidirectional residual fusion is proposed. On one hand, a multi-feature selection module based on mixed attention is designed to select multiple features of underwater images. Furthermore, the bidirectional residual fusion is used to optimize traditional U-shaped skip connection, which realizes high-efficiency expression of image features and effectively restores the texture and color of underwater images. In addition, mixed attention is introduced and content-aware loss and style-aware loss are proposed in the discriminator to ensure that the enhanced image is consistent with the clear image in terms of global content, local texture, and style features. The PSNR of the proposed model is improved by 6% and 2%, respectively, compared with the existing unsupervised and supervised models. Additionally, SSIM is improved by 4% and 3%, respectively. With a significant enhancement effect on underwater images, the proposed method demonstrates superiority over other existing methods in terms of color fidelity and saturation.

Keywords: unsupervised models; recurrent generative adversarial networks; multi-feature selection; bidirectional residual fusion; underwater image enhancement

0 引言

海洋蕴含着十分丰富的可利用资源,这些资源的有效利用对人类社会发展有着重要影响。由于水下环境恶劣复杂,需要水下机器人替代人类在危险的水下环境中进行探测,得到的高质量水下图像与水下视频可以提供探测者丰富的海洋信息,并有利于在海洋中完成定位、目标识别和路径规划引导等任务。光在水下环境传播过程中会发生衰减和散射,水下图像通常会出现能见度低、颜色偏差、失真、对比度低等问题^[1-2]。波长较短的蓝绿色光在水下环境中衰减较慢,故大部分的水下图像呈现蓝绿色调^[3-4]。这些问题影响了水下图像的质量,进而对水下目标检测、识别等任务造成影响。因此,如何将低质量的水下图像增强到高质量图像已成为水下实际应用中需要解决的首要问题。

严浙平等^[5]将现有的水下增强方法分为非物理模型、物理模型和深度学习方法。如直方图均衡化优化^[6]、伽马函数校正^[7]、自适应灰色世界方法^[8]、白平衡^[9]、基于融合的方法^[10]等非物理模型方法,只直接调整原始图像的像素值来提高视觉效果,忽略了水下成像特性,限制了它们的增强效果。而物理模型的方法主要创建水下图像退化过程中的物理模型,利用该模型对图像退化过程进行反演得到清晰图像。常见的有水下暗通道先验(underwater dark channel prior,UDCP)^[11]、自适应颜色校正^[12-13]、水下光衰减先验(underwater light attenuation priori,ULAP)^[14]、红通道补偿^[15]等。但单一的物理参数在多变的水体环境下对图像的颜色校正会产生严重偏差,鲁棒性较弱。

由于深度学习在各个领域都优于传统方法,故卷积神经网络(convolutional neural networks,CNN)^[16-18]和生成对抗网络(generative adversarial networks,GANs)^[19-21]依靠着强大的表征学习能力也被应用于水下图像增强领域,并迅速发展。Chen等^[22]提出了一种包含对抗和临界多分支判别器的生成对抗网络的增强方案,该方案在保持图像内容的同时去除水下噪声,能实时并自适应地改善水下浑浊图像质量。Guo等^[23]将多尺度、残差网络和密集连接高效组合起来,消除原始水下图像的色彩畸变并保留图像的细节信息与图像特征。Li等^[24]提出了一种基于先验的卷积神经网络来增强水下图像,该模型模拟多种水下浑浊图像,联合优化多项损失输出清晰的水下图像,无需对物理模型的参数进行估计。Wang等^[25]提出一种双色空间的水下图像增强卷积神经网络,该网络有效结合RGB与HSV两种颜色空间融入到CNN中,注意力通道块结合RGB和HSV两者输出图像的优点对每个像素重新分配权重输出最终图像。Ren等^[26]在U-

Net中添加Swin Transformer^[27]来捕获全局特征信息,并融合卷积与核心注意力以增强Transformer块对局部特征的的关注能力,该方法能有效去除图像色彩伪影。Liu等^[28]提出一种并行注意力与自适应学习模块,结合不同的空间信息以解决水下图像色偏与低光照度的问题。

由于真实的水下环境与模拟设置的水下模型参数差异性较大,上述利用水下先验信息合成的成对数据集进行训练的有监督模型,在实际应用中增强的水下图像会出现颜色特征丢失、色彩真实度差等问题,模型泛化能力弱,影响水下设备后续处理任务的精度。因此无需依赖成对数据集的无监督学习是水下图像增强的一个新研究方向。Islam等^[29]提出了一种能对水下图像的整体结构、颜色与风格特征进行评估感知的条件生成对抗网络模型,该模型能使用成对或非成对的水下图像数据集训练模型来提高水下图像的视觉效果。Lu等^[30]提出一种多尺度循环生成网络的方法,将循环生成对抗网络^[31]与结构相似性指数函数融合进而改善水下图像质量,使图像从水下浑浊风格到转换到清晰风格,实现了无监督的水下图像增强。Han等^[32]联合对比学习与生成对抗网络来最大化浑浊图像与清晰图像之间的互信息,重构出清晰的水下图像,实现一种图像到图像的无监督转换框架。Park等^[33]采用一种自适应加权的判别器循环一致性网络模型,使用自适应加权法制约内容判别器与风格判别器之间的损失,重建出较好的水下图像。以上无监督方法虽然解决了网络模型对配对数据集的过度依赖问题,但它们增强过程中容易丢失多种特征信息,导致去除色偏不干净且易出现红色伪影,在细节信息恢复以及色彩校正上仍需进一步加强。

为了使深度学习模型能摆脱成对数据集依赖的同时,有效解决浑浊水下图像的异常色彩和细节模糊等问题,本文提出了一种多特征选择与双向残差融合的无监督生成对抗网络(multi-feature selection and bidirectional residual fusion unsupervised generation adversarial networks,MSBFU-GAN)学习水下图像中浑浊域和清晰域的不成对映射关系。该模型既能改善因水体导致的畸变色域,又能补足纹理细节特征并恢复图像整体色彩,使处理后的图像更加的鲜艳自然。主要贡献如下:

- 1) MSBFU-GAN 构建了由不对称卷积、不同膨胀率的平滑空洞卷积和混合注意力组成的多特征选择模块,对水下图像不同尺度的特征进行优势选择。

- 2) 针对特征信息易丢失的问题,MSBFU-GAN 引入双向残差融合的方式逐渐融合图像多感受野和全局特征信息,保留水下图像复杂多变的特征。

- 3) MSBFU-GAN 的混合注意力判别器,学习输入图像的上下文信息,更好地捕获像素之间的依赖,提高对图像特征的判别能力,以激励生成器产生更自然的图像。

4) 本文设计了使用 RepVGG 预训练的内容感知损失与风格感知损失,通过感知图像间全局内容、局部细节和风格特征的差异来训练模型,以校正图像视觉效果并提高生成器生成的图像质量。

1 MSBFU-GAN 的无监督水下图像增强方法

本文设置浑浊域 I 是水下未处理的图像集合,清晰域 J 是水下清晰目标图像集合。整体框架如图 1(a) 所示,MSBFU-GAN 由两个生成器和两个判别器构成清晰方向和浑浊方向网络,共同学习不成对浑浊域 I 和清晰域 J 之间的转化。生成器 G 和生成器 H 分别学习 $I \sim J$ 的映射 ($F_1: I \rightarrow J$) 和 $J \sim I$ 的映射 ($F_2: J \rightarrow I$)。图 1

(b) 以清晰方向为例。浑浊图像 i 经过生成器 G 生成伪清晰域图像 $G(i)$,为防止 I 的所有浑浊图像都被映射成 J 中的某一张清晰图像,再经过生成器 H 生成重构浑浊域图像 $H(G(i))$,以保证循环一致性。同时,判别器 $D(i)$ 与 $D(j)$ 用于区分 $G(i)$ 和 $H(j)$ 与对应的真实域样本,推动生成器与自身的更新,直到模型达到纳什平衡。浑浊方向过程与以上相同。生成器和判别器详细内容将在第 1.1 和 1.2 节进行介绍。

本文除了使用对抗损失和循环一致损失完成无监督生成对抗网络的基本运行,还添加了内容感知损失函数和风格感知损失函数,保障在生成的图像中保留足够的内容和风格特征信息。具体介绍将在 1.3 节进行。

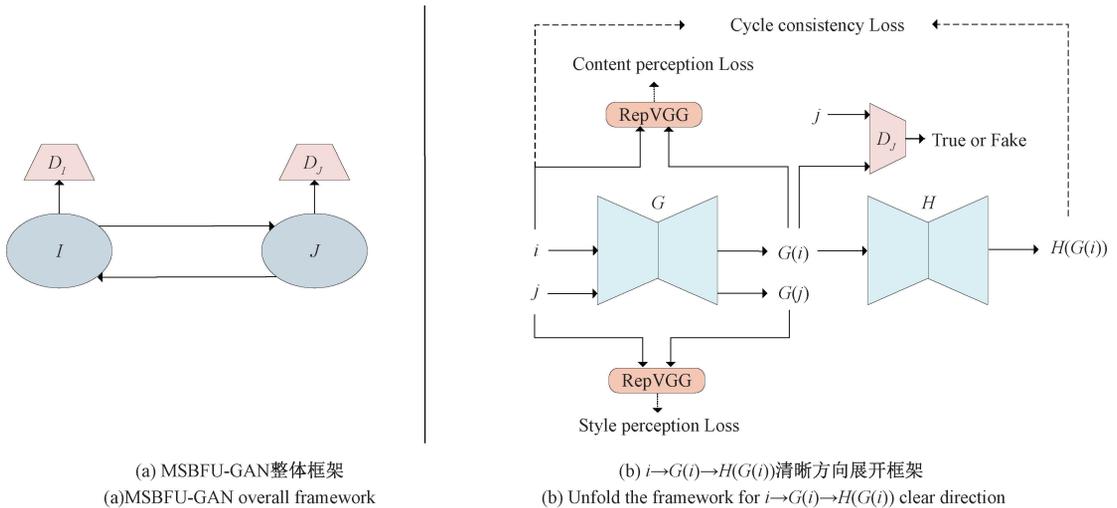


图 1 本文无监督模型框架

Fig. 1 Unsupervised model framework of this paper

1.1 MSBFU-GAN 生成器网络结构

MSBFU-GAN 生成器结构如图 2 所示,共 4 个部分组成:U 型架构、多特征选择模块 (multi-feature selection module, MFS)、全局特征抽取模块 (global feature extraction module, GFE)、双向残差融合 (bidirectional residual fusion, BRF)。基础 U 型网络包含 4 个编码级与解码级,两个 CIR 单元 (Conv 层+InstanceNorm 层+ReLU 层),4 个下采样与上采样。其中,使用多残差块 (multi-residue blocks, MRB) (图 3(a)) 作为 MSBFU-GAN 生成器的 4 个编码-解码级,以增加网络深度,使网络层能够学习到更加复杂度表达。生成器的下采样使用 2×2 卷积,避免使用池化下采样带来的特征丢失。上采样时采用双线性插值法,双线性插值法比反卷积的优势在于效率高,无需训练就能利用数学公式计算出周围像素色彩丢失的像素色彩,同时使图像干净平滑。

为在处理图像特征信息时能灵活选择关注不同类型的信息,更好地解决水下复杂环境对图像区域的影响,本文采用通道注意力 (channel attention, CA) 和空间注意力 (spatial attention, SA) [34] 综合通道和空间的优势信息。混合注意力模块 (mixed attention module, MA) 如图 3(c) 所示。

压缩激励网络 (squeeze-and-excitation networks, SENet) [35] 与频率通道注意力网络 (frequency channel attention networks, FcaNet) [36] 等通道注意力通常使用图 4 所示的非线性全连接层来捕捉跨通道信息,但这种方式会破坏原始特征图的结构,导致空间信息丢失。故本文 CA 用卷积核为 m 的 1 维卷积将 m 个邻近特征通道相关联,以捕捉局部跨通道交互信息。这种方法有益于更精细的特征表示,使模型更适合处理光照不均匀的水下图像。其中, m 表示局部通道交互的有效范畴。为了避免在训练和调用过程中手动调整 m 的大小,本文采用了高

1) 混合注意力模块

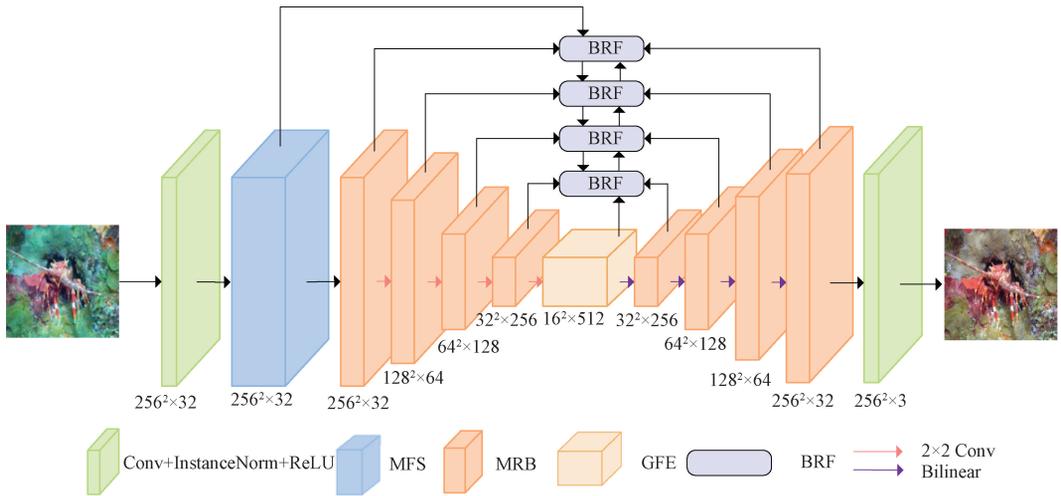
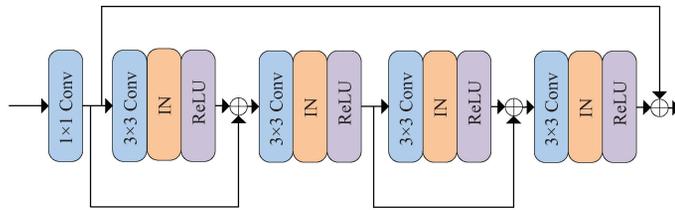


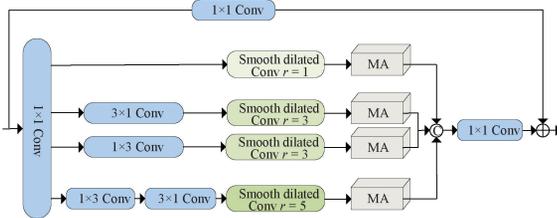
图2 MSBFU-GAN 生成器整体结构

Fig. 2 MSBFU-GAN Overall structure of generator



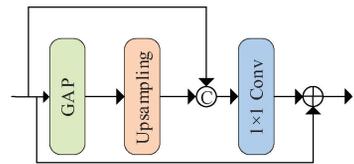
(a) 多残差块

(a) Multi-residue blocks(MRB)



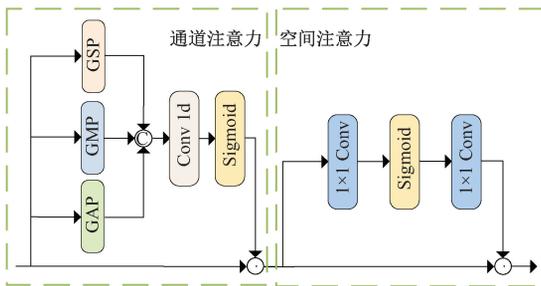
(b) 多特征选择模块

(b) Multi-feature selection module(MFS)



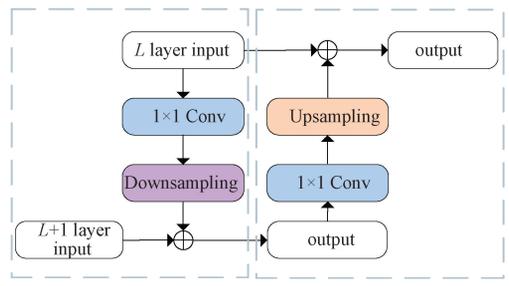
(d) 全局特征抽取模块

(d) Global feature extraction module(GFE)



(c) 混合注意力模块

(c) Mixed attention module(MA)



(e) 双向残差融合

(e) Bidirectional residual fusion(BRF)

图3 MSBFU-GAN 生成器关键组件结构

Fig. 3 Critical component structure of MSBFU-GAN generator

效通道注意力网络(EcaNet)^[37]的方法,自适应地确定 m 的大小。表达式为:

$$m = f_m(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{\alpha}{\gamma} \right\rfloor_{\text{odd}} \quad (1)$$

式中: $\lfloor \lambda \rfloor_{\text{odd}}$ 为与 λ 接近的奇数, C 为通道数, γ 数值为 2, α 数值为 1。CA 除常用的全局平均池化外,还添加了全局最大和标准差池化,以增加表达能力。输入特征经过全局池化将通道方向的信息进行压缩,并对其进行拼接送入 1 维卷积和 sigmoid 函数,挑选相对重要的特征纳入特征图。CA 的表达式如下:

$$Y_{CA} = \sigma(C_m^{1D}(F_C[GAP(T_C), GMP(T_C), GSP(T_C)])) \cdot T_C \quad (2)$$

式中: $Y_{CA} \in R^{8 \times 256 \times 256}$ 为 CA 的输出,其输入是 $T_C \in R^{8 \times 256 \times 256}$,表示上阶段输出第 C 通道的特征。 GAP 、 GMP 和 GSP 分别是平均池化、最大池化和标准差池化操作。 C_m^{1D} 为卷积核为 m 的 1 维卷积。 F_C 表示拼接操作,特征图按通道维度进行拼接。 σ 是 sigmoid 函数,对输入 T_C 的通道维度实行特征选择操作。

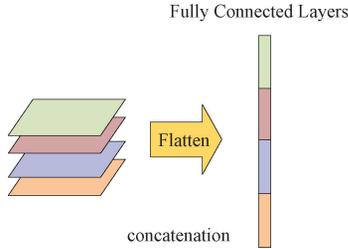


图 4 全连接层展平操作

Fig. 4 Fully connected layer flattening operation

空间注意力 (coordinate attention, SA) 用以校对特征图在空间上的关键位置,同时忽略相对不相关的位置,有助于恢复水下图像纹理细节等高频信息。SA 通过 1×1 的卷积将 CA 输出特征图的通道维度压缩至 1,并将压缩后的结果传递给 sigmoid 函数得到空间注意力特征映射,最后对空间注意力特征图和输入特征图进行元素乘法生成最终需要的 MA 特征图。其中, 1×1 卷积并不改变特征图大小,在空间方向上保证了特征的一致性。故 MA 的表达式如下:

$$\begin{aligned} Y'_{MA} &= \sigma(\text{Conv}(Y_{CA})) \\ Y_{MA} &= \text{Conv}(Y'_{MA}) \cdot Y_{CA} \end{aligned} \quad (3)$$

式中:输出 $Y'_{MA} \in R^{1 \times 256 \times 256}$ 为空间注意力操作, Conv 为 1×1 卷积,用于扩充或压缩通道维度,则 MA 模块的输出 $Y_{MA} \in R^{8 \times 256 \times 256}$ 。

2) 多特征选择模块

本文提出 MFS 采集各图像的不同优势特征,解决单一卷积层的感受野无法提取不同大小目标的特征和不同优势信息的问题,使网络能更好处理由水中环境导致的色偏、模糊等畸变区域。如图 3(b) 所示,与传统的多径结构不同, MFS 使用不同膨胀率 ($r=1, 3, 5$) 的平滑膨胀

卷积^[38]代替普通的卷积,以此获得 CIR 单元输出特征图不同层次的特征区域信息, MA 分别对各平滑膨胀卷积得到的不同特征进行选择,选择出的不同特征将拼接到一起,最后经过卷积层自适应获得优势特征,并与输入累加输向 BRf。MFS 表达式如下:

$$\begin{aligned} O_i &= Y_{MA}(f_{SCov}^i(x)) \\ Y_{MFS} &= \text{Conv}(F_C[O_0, O_1, O_2]) + \text{Conv}(x) \end{aligned} \quad (4)$$

式中: MFS 的输入为 x , i 和 f_{SCov}^i 分别表示第 i 条路径和第 i 条路径的不对称卷积与平滑膨胀卷积的组合函数, $i \in [0, 3]$ 。 O_i 是第 i 条路径的输出特征, MFS 中的维度变化为 $x \in R^{32 \times 256 \times 256}$, $O_i \in R^{8 \times 256 \times 256}$, $Y_{MFS} \in R^{32 \times 256 \times 256}$ 。

3) 全局特征抽取模块

普通卷积的卷积核只能提取图像的局部特征,而模型需要兼顾图像的全局特性来解决浑浊域水下图像的低光照、颜色失真等问题。因此 MSBFU-GAN 引入一个如图 3(d) 所示的 GFE,使模型能够采集水下图像的全局特征,以改善图像的整体色彩饱和度与亮度。GFE 表达式如下:

$$\begin{aligned} Y'_{GFE} &= F_{up}(\text{GAP}(Y_{CRB,4}^n)) \\ Y_{GFE} &= \text{Conv}(F_C[Y'_{GFE}, Y_{CRB,4}^n]) + Y_{CRB,4}^n \end{aligned} \quad (5)$$

式中:第 4 个编码级的输出 $Y_{CRB,4}^n \in R^{512 \times 16 \times 16}$ 作为 GFE 的输入, F_{up} 为上采样操作,将池化后的特征图恢复到与输入相同的尺寸,则 GFE 中 $Y'_{GFE} \in R^{512 \times 16 \times 16}$, $Y_{GFE} \in R^{512 \times 16 \times 16}$ 。

4) 双向残差融合

解码器单独使用上采样易使输出图像模糊,且在深层特征图中可能会丢失许多与图像轮廓和纹理相关联的低级特征。如图 3(e) 所示, MSBFU-GAN 使用 BRf 将 MFS 选择的优势特征与 GFE 抽取的全局特征逐级融合到解码器部分,不仅能弥补编码级和解码级之间的语义鸿沟,且让获得的多种特征高效表达,提供相对高质量的特征传递到解码部分。BRf 的输入来自模型的浅层特征 (L 层) 和相应深层特征 ($L+1$ 层), 整数 $L \in [1, 4]$ 。BRf 表达式如式 (6) 所示:

$$\begin{aligned} Y'_{BRF} &= F_{down}[\text{Conv}(X_L)] + X_{L+1} \\ Y_{BRF} &= F_{up}[\text{Conv}(Y'_{BRF})] + X_L \end{aligned} \quad (6)$$

其中, $X_L \in R^{C \times HW}$ 来自 L 层 BRf 的输出特征,仅当 $L=1$ 时,为 MFS 的输出。 $X_{L+1} \in R^{2C \times \overline{HW}}$ 是来自 $L+1$ 层 MRB 的输出特征,仅当 $L=4$ 时, X_{L+1} 为 GFE 的输出, \overline{HW} 表示 H 和 W 大小的 $1/2$ 。 F_{down} 为下采样操作,缩小特征图大小。 F_{up} 为上采样操作,利用双线性插值法扩充特征图。 Conv 为 1×1 卷积,将通道扩充至 $2C$ 或缩减至 C ,使 $Y'_{BRF} \in R^{2C \times \overline{HW}}$ 和 $Y_{BRF} \in R^{C \times HW}$ 分别跟 X_{L+1} 和 X_L 保持一致的特征维度。

1.2 MSBFU-GAN 判别器网络结构

马尔代夫判别器(PatchGAN)^[39]是一种基于补丁级区域信息对图像进行鉴别的判别器,能高效捕捉水下生物的边缘纹理等高频特征。本文在 $16 \times 16 \times 1$ 的 PatchGAN 中引入混合注意力(MA),以训练判别器更好

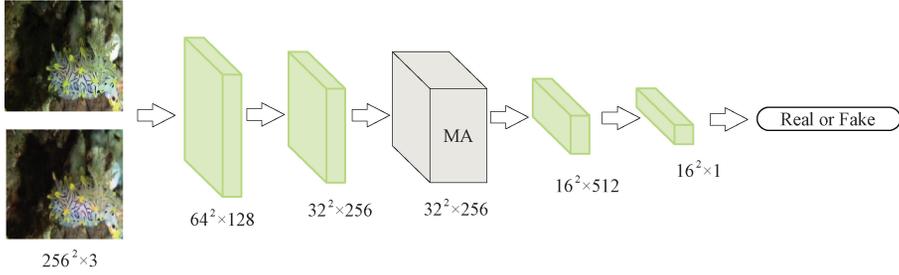


图5 MA 判别器

Fig. 5 MA discriminator

1.3 损失函数

1) 对抗损失

对抗损失是一种优化生成器和判别器相互迭代的损失函数,它约束生成图像 $G(i)$ 和 $H(j)$ 分别与清晰图像 j 和浑浊图像 i 的分布相同。对抗损失可表示为:

$$L_{adv}(G, D_j) = E_{j \sim P_{data}(j)} [(D_j(j))^2] + E_{i \sim P_{data}(i)} [(1 - D_j(G(i)))^2] \quad (7)$$

$$L_{adv}(H, D_i) = E_{i \sim P_{data}(i)} [(D_i(i))^2] + E_{j \sim P_{data}(j)} [(1 - D_i(H(j)))^2] \quad (8)$$

式中: $P_{data}(i)$ 和 $P_{data}(j)$ 分别表示水下浑浊域 I 中图像 i 和清晰域 J 中图像 j 的分布, $E_{i \sim P_{data}(i)}$ 和 $E_{j \sim P_{data}(j)}$ 分别表示 i 服从 $P_{data}(i)$ 条件下的均值和 j 服从 $P_{data}(j)$ 条件下的均值。

2) 循环一致损失

循环一致损失是 $H(G(i))$ 与 i , $G(H(j))$ 与 j 之间的损失函数,它保证了生成图像与对应输入图像的内容一致,对模型进行监督。循环一致损失可表示为:

$$L_{cyc}(G, H) = E_{i \sim P_{data}(i)} [\|H(G(i)) - i\|_1] + E_{j \sim P_{data}(j)} [\|G(H(j)) - j\|_1] \quad (9)$$

式中: $\|\cdot\|_1$ 表示 L_1 范数。

3) 内容感知损失

本文希望输入图像的细节信息保持不变的同时,又能校准输入图像的颜色以提高生成图像的视觉质量。为此,引入 RepVGG^[40] 预训练网络的 Conv3_2 层提取输入图像 i j 和生成图像 $G(i)$ 、 $H(j)$ 的特征图,通过逐元素减法让其视觉上的差异最小化。同时,采用平滑 L_1 范数^[41] 约束输入图像和生成图像的像素,以提高网络鲁棒性和收敛性,更好地恢复水下图像的颜色特征。内容感知损失公式如下:

地鉴别生成的水下图像,如图5所示。MA判别器从图像中获取特征,然后通过多次卷积和混合注意力机制生成一维输出,进而判断获取的特征是否属于特定的类别,有效捕获图像水体环境和生物特征信息,增强判别器的性能。

$$L_{repVGG}(G, H) = E_{i \sim P_{data}(i)} [\|\psi(G(i)) - \psi(i)\|_{s_1}] + E_{j \sim P_{data}(j)} [\|\psi(H(j)) - \psi(j)\|_{s_1}] \quad (10)$$

式中: $\|\cdot\|_{s_1}$ 为平滑 L_1 范数, $\psi(\cdot)$ 表示预训练 RepVGG 网络提取的高级特征。RepVGG 采用多分支结构进行训练,使用结构重参数化单分支网络进行推理,使其提取图像特征更加高效,有利于获得更加精细的像素信息,增强图像内容颜色更加生动。

4) 风格感知损失

添加 RepVGG 预训练网络和平滑 L_1 范数的风格感知损失用于测量生成图像与输入图像特征表示之间的距离,控制增强图像整体色调倾向自然,保证增强图像和预期的目标图像的风格一致。风格感知损失可表示为:

$$L_{repsty}(G, H) = E_{j \sim P_{data}(j)} [\|\psi(G(j)) - \psi(j)\|_{s_1}] + E_{i \sim P_{data}(i)} [\|\psi(H(i)) - \psi(i)\|_{s_1}] \quad (11)$$

5) 总体损失

因此,整个训练模型的总损失为:

$$L_{total} = L_{adv}(G, D_j) + L_{adv}(H, D_i) + \alpha L_{cyc}(G, H) + L_{repVGG}(G, H) + \beta L_{repsty}(G, H) \quad (12)$$

式中: α 和 β 分别对应循环一致损失、风格感知损失的权重系数。其中 $\alpha = 10$, $\beta = 5$ 。

2 实验结果与分析

在本章中,设计了一些对比实验和消融实验来验证 MSBFU-GAN 的优越性与有效性。首先,将其他先进方法分为有监督组 (UWCNN^[24]、Shallow-UWnet^[17]、UIEC²-Net^[25]、URSCT-SESR^[26]) 和无监督组 (CycleGAN^[31]、FunieGAN^[29]、CWR^[42]、F-LSesim^[43]) 两组模型,它们分别与 MSBFU-GAN 的增强结果进行定性和定量对比分析。其中,对每组模型都设置相同参数,并分别使用成对

或不成对数据集上进行训练,以保证对比实验的公平性。此外,通过对 MSBFU-GAN 的重要结构损失函数进行消融实验和评估分析,进一步验证了其有效性。

2.1 数据集设置

本文使用 Islam^[29] 提供的 EUVP 的数据集用于本文模型、对比实验模型以及消融实验模型的训练和评估。其中,有监督组在 EUVP 的 11 435 张成对训练样本、1 270 张验证样本,515 张测试样本上进行训练与测试。无监督组采用 EUVP 不成对数据集,但此部分数据样本相对较少,用于网络训练可能会出现过拟合问题,故本文使用旋转、亮度变换、翻转等扩充操作增加数据样本的多样性,避免训练过拟合的问题。扩充操作后的不成对数据集有 13 626 张浑浊域图像和 9 420 张清晰域图像作为训练样本用于训练,并使用其中默认划分的 330 张验证样本与 515 张测试样本用于评估。

2.2 实验设置

本文使用 Pycharm 平台进行代码编写,选用 Python3.8 编程语言在 PyTorch1.10.0 深度学习框架下完成模型的搭建与实验,所有实验均在 NVIDIA GeForce GTX 3090 和运行内存 24 GB GPU 内存服务器上运行,操作系统为 Ubuntu-20.04。训练过程中,随机截取 256×256 大小的区域作为模型输入,批量大小为 4,采用 Adam 算法对网络进行优化,初始学习率设为 0.000 3,迭代 10 轮后学习率开始衰减,网络总共迭代 150 轮。

2.3 定量评价指标

为了能用客观数据量化无监督模型的性能来验证 MSBFU-GAN 有效性,本文不仅选择常用的峰值信噪比 (peak signal to noise ratio, PSNR)、结构相似性 (structural similarity index measure, SSIM)^[44] 作为评价指标,并且采用无需参考图像对比的水下图像质量测量 (underwater image quality measurement, UIQM)^[45] 对两组各模型的图像质量进行客观评价与定量分析。PSNR 是峰值信号能量与噪声能量均值比,它量化了图像的失真程度,其值越高,证明增强图像失真越少,真实性越高;SSIM 量化了增强图像的结构信息完整程度,其值反映增强图像与参考图像结构相似度的指标,它主要用来评价图像的光照、对比度以及结构 3 个方面^[46]。它的值越大,表示两者之间结构更加相似,更符合人眼的直观感受;UIQM 由水下图像色彩测量 UICM、水下图像清晰度测量 UISM、水下图像对比度测量 UIConM 组成,量化了水下图像的整体质量,能反映出处理后的真实水下图像增强效果。它值越大,增强图像在色彩平衡、清晰度、对比度上表现越好。

2.4 对比实验

1) 无监督组评估分析

图 6 展示了近几年 4 种无监督模型和 MSBFU-GAN

在 EUVP 不成对数据集中的水下图像增强效果。其他无监督模型增强后的图像均存在蓝绿色偏去除不干净的问题,视觉效果较差。其中,CycleGAN 处理的水下图像细节模糊,整体亮度相对偏暗,且部分图像与 CWR 的部分图像都出现红色伪影的问题。MSBFU-GAN 不受水体环境影响,不仅去除了由水体原因导致的蓝绿色偏、红色伪影和浑浊模糊现象,而且色彩效果更加自然生动,在视觉上更接近参考图像。

MSBFU-GAN 与其他无监督模型的定量比较如表 1 所示,参数加粗为最优值。表 1 清晰地反映出,与其他增强模型相比,本文模型在 PSNR、SSIM 和 UIQM 指标上获得了最优,MSBFU-GAN 在无监督组实验中的 PSNR 达到 27.439 dB,比其他模型至少高出 1.58 dB。由此证明 MSBFU-GAN 失真少,拥有更高的对比度和更多的细节特征。

表 1 无监督组的定量比较

Table 1 Quantitative comparison of unsupervised group

模型	PSNR	SSIM	UIQM
FunieGAN	25.455±2.658	0.789±0.077	2.845±0.35
CycleGAN	25.859±2.964	0.803±0.071	2.838±0.41
CWR	24.952±2.975	0.8±0.067	2.85±0.43
F-LSeSim	25.2±2.832	0.79±0.066	2.885±0.425
MSBFU-GAN	27.439±2.71	0.836±0.065	2.89±0.4

2) 有监督组评估分析

图 7 为 MSBFU-GAN 与 4 种有监督模型的效果对比图。其他有监督模型植被色彩暗淡,与原始浑浊域图像明显不符。其中,UWCNN 在一定程度上能纠正色偏,但增强图像相对模糊;Shallow-UWnet 生成的图像也存在模糊问题,且饱和度过低,部分颜色特征丢失。UIEC²-Net 对消除蓝色色偏无明显效果,且增强后的图像对比度过高,视觉效果不自然。URSCT-SESR 生成的图像整体亮度有限,颜色较暗。

在表 2 中,MSBFU-GAN 的 PSNR 和 SSIM 指标同样均优于其他有监督模型,但 UIQM 值仅次于 UIEC²-Net。UIQM 主要关注水下图像的色彩、清晰度以及对比度,并没有真实值作为参考,容易忽略原始水下物体的真实色彩。然而 UIEC²-Net 处理后的图像仍存在大量的蓝色色偏,但其本身有着良好的清晰度和对比度,所以 UIQM 值相对较高。结合以上定性定量评价可以得出,MSBFU-GAN 在饱和度和亮度上均得到了不小提升,能有效还原水下图像色彩真实度,相比原始浑浊域中的水下图像提升较大。

综上所述,本文提出的 MSBFU-GAN 无论是在成对数据集还是不成对数据集上,均远胜于其他模型的增强效果,具有很好的视觉效果和定量指标,证明了本文所提

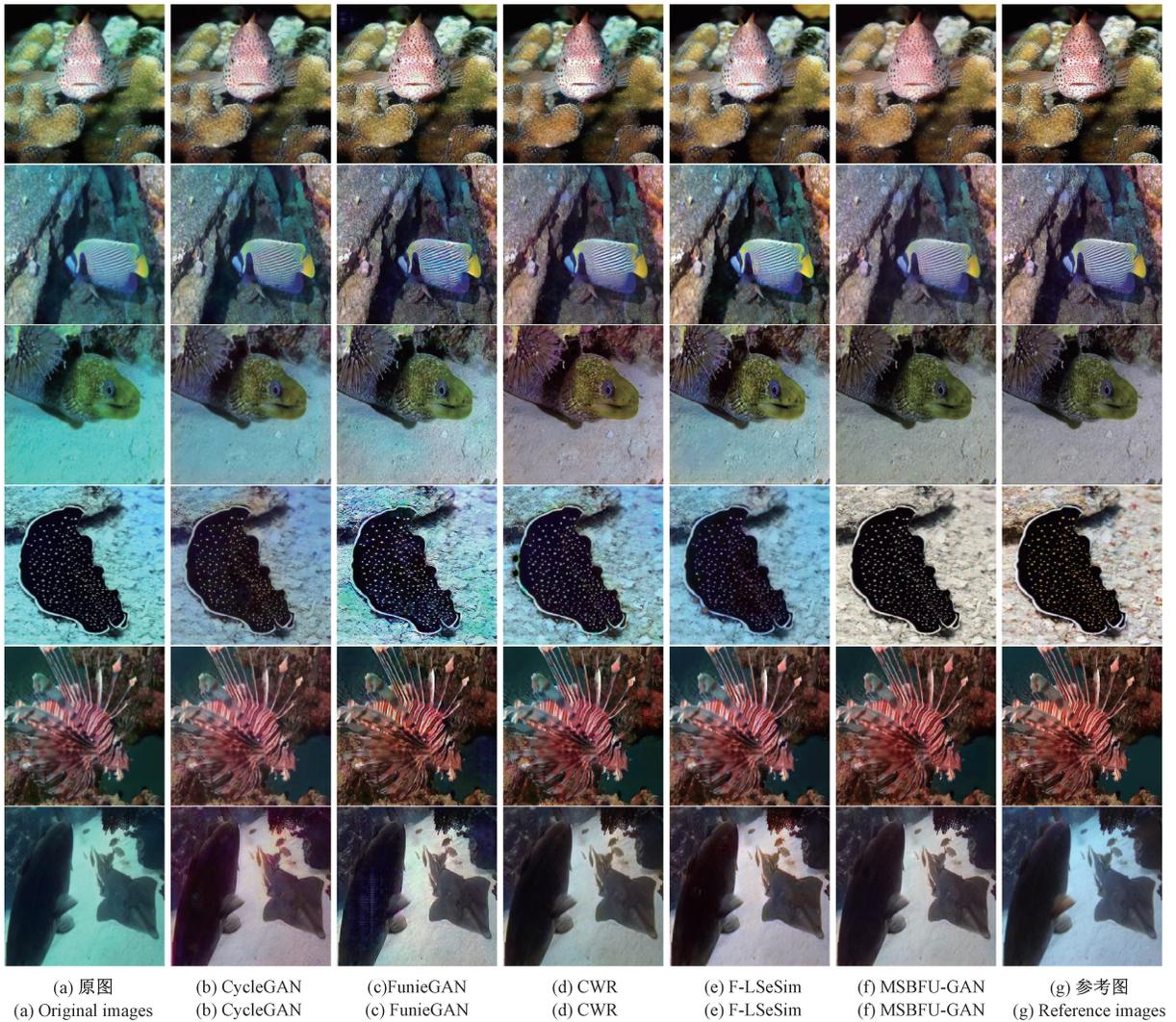


图 6 无监督组模型的视觉效果对比

Fig. 6 Visual effects comparison of unsupervised group models

出的无监督模型的优越性。

表 2 有监督组模型的定量比较

Table 2 Quantitative comparison of supervised group

模型	PSNR	SSIM	UIQM
UWCNN	25.42±3.14	0.805±0.062	2.842±0.382
Shallow-UWnet	26.762±2.978	0.81±0.066	2.864±0.327
UIEC ² -Net	21.597±3.65	0.784±0.087	2.999±0.379
URSCT-SESr	27.012±2.687	0.827±0.061	2.855±0.372
MSBFU-GAN	27.42±2.838	0.839±0.061	2.868±0.423

2.5 消融实验

为了验证 MSBFU-GAN 各个部分的必要性和有效性,本文分别对其重要结构和损失函数进行消融实验。

1) 关键组件消融实验

图 8 和表 3 显示了消除 MFS、GFE、BRF 而使用传统 U 型跳跃连接、消除判别器中的混合注意力消融实验结果。其中,表 3 参数加粗为最优值。结合两者可以看出, MFS 能校正图像的颜色,改善图像对比度并提升纹理细节特征,去除 MFS 后 PSNR 指标下降了 1.585 dB, SSIM 指标也有所降低,说明了 MFS 在模型中的合理性。增加 GFE 使网络关注浑浊域图像整体特征,从而改善图像亮度和色彩饱和度,提升各项指标。在无 BRF 的情况下使用 MFS 和 GFE 的效果较差,由于丢失了图像的多种特征,导致模型未能有效学习到图像的水体信息,不善于处理纹理密集、光照不均匀等复杂水下图像。使用 MA 判别器的模型能更好捕捉到水下图像的生物特征,使增强图像在局部细节上更加自然,整体风格更贴近清晰域图像的风格。

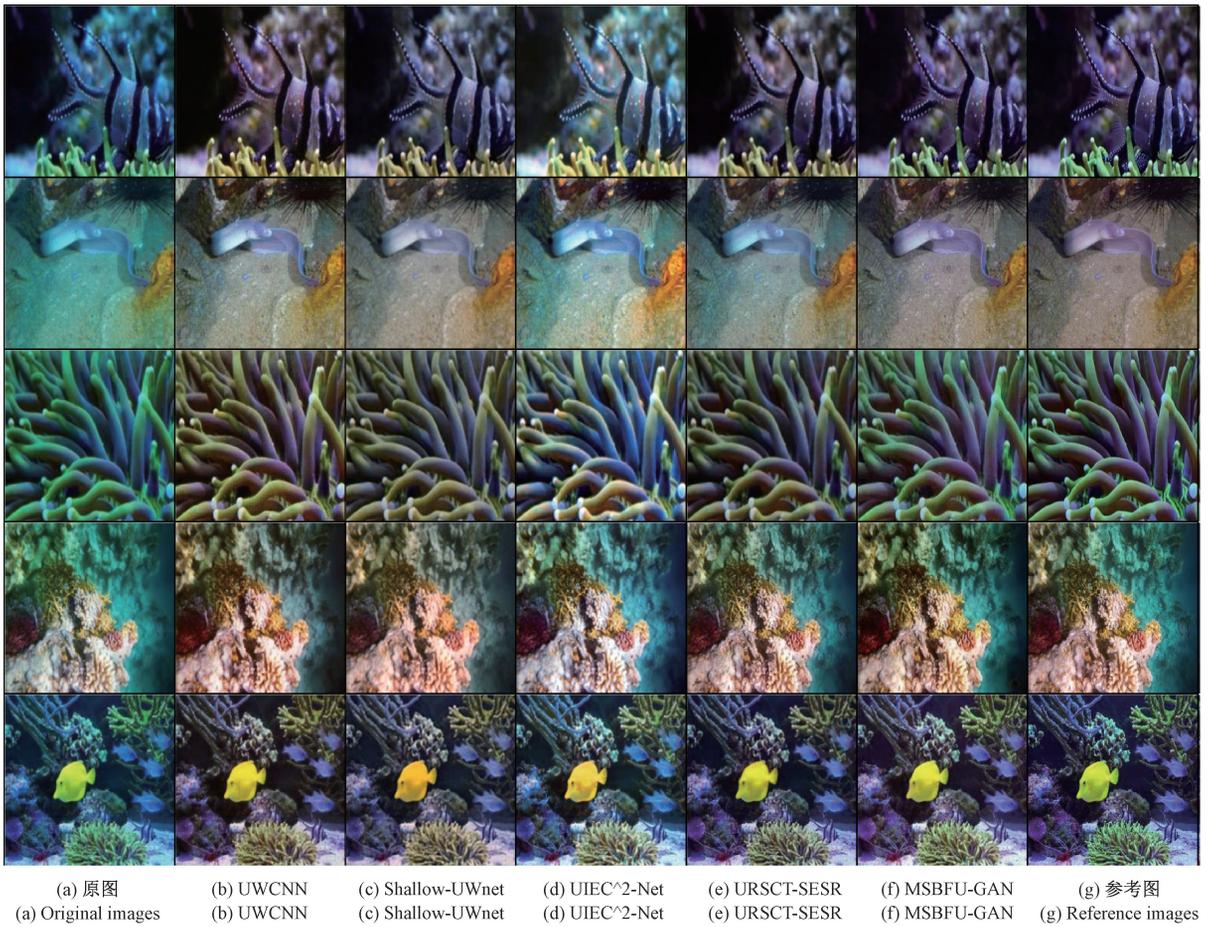


图 7 有监督组模型的视觉效果对比

Fig. 7 Visual effects comparison of supervised group models

表 3 MSBFU-GAN 关键组件消融实验的定量比较
Table 3 Quantitative comparison of MSBFU-GAN critical component ablation experiments

模型	PSNR	SSIM
w/o ⁽¹⁾ MFS	25.854±2.433	0.818±0.06
w/o GFE	26.711±2.896	0.827±0.068
w/o BRF	26.984±2.784	0.834±0.064
w/o MD ⁽²⁾	27.22±2.748	0.836±0.062
MSBFU-GAN	27.439±2.71	0.836±0.065

注:(1) w/o 表示 without; (2) MD 是 MA Discriminator 缩写。

2) 关键损失函数消融实验

图 9 和表 4 显示了消除内容感知损失、风格感知损失的消融结果。其中,表 4 参数加粗为最优值。定性和定量两方面均展示了内容感知损失、风格感知损失在模型中的重要性。其中,风格感知损失有效解决图像增强后饱和度和亮度过高问题,网络性能得到了有效的提升。内容感知损失的增加显著改善了水下图像的颜色特征,使图像鲜艳饱满。各损失函数均有利于 MSBFU-GAN 提

升视觉效果,尤其去除风格感知损失,MSBFU-GAN 的 PSNR 和 SSIM 指标将被最小化。

表 4 MSBFU-GAN 损失函数消融实验的定量比较
Table 4 Quantitative comparison of MSBFU-GAN loss function ablation experiments

模型	PSNR	SSIM
w/o rep_sty	23.588±2.951	0.758±0.074
w/o ⁽¹⁾ repVGG	26.032±2.572	0.828±0.064
MSBFU-GAN	27.439±2.71	0.836±0.065

注:w/o 表示 without。

以上实验结果表明,完整结构的 MSBFU-GAN 拥有最佳性能,所提的无监督模型结构对水下图像增强任务有着非常显著的效果。

3 结 论

本文提出一种多特征选择与双向残差融合的无监督

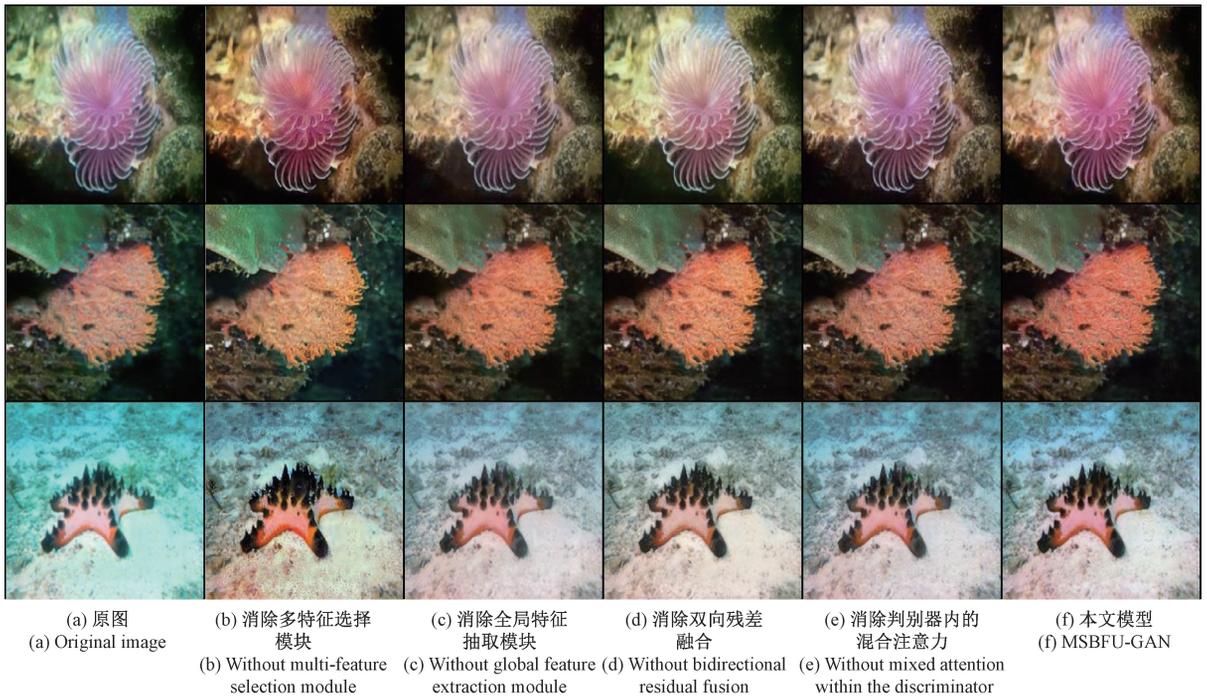


图 8 MSBFU-GAN 关键组件消融实验的定性结果

Fig. 8 Qualitative results of MSBFU-GAN Critical component ablation experiment

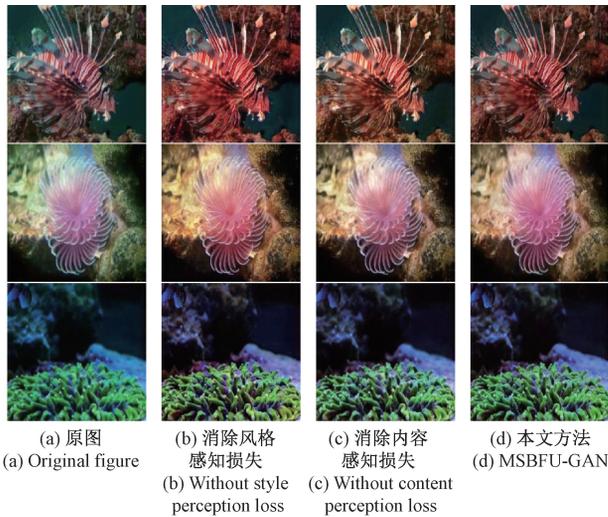


图 9 MSBFU-GAN 关键损失函数消融实验定性结果

Fig. 9 Qualitative results of MSBFU-GAN critical loss function ablation experiment

水下图像增强方法—MSBFU-GAN。本文生成器采用 BRF 取代传统的跳跃连接,并配合 GFE 和 MFS,兼顾水下图像的全局和局部畸变信息,高效遏制无用特征的表达。MA 判别器具有出色的生物特性捕捉能力,能激励生成器产生更自然的水下图像。此外,内容感知损失与风格感知损失的使用,能最小化输入图像与生成图像的

特征差异,改善结果图像的视觉效果。消融实验验证了 MSBFU-GAN 重要组成部分的有效性,对比实验证明了 MSBFU-GAN 效果优于其他无监督和有监督模型,能较好的改善水下成像造成的色偏、伪影和细节模糊等问题。

未来工作中,将尝试在水下机器人上部署增强模型,并探索如何在保证高性能的前提下对模型轻量化,以便水下机器人对图像能进行实时增强。

参考文献

[1] LI Y, ZHANG Y, XU X, et al. Dust removal from high turbid underwater images using convolutional neural networks[J]. Optics & Laser Technology, 2019, 110: 2-6.

[2] FABBRI C, ISLAM M J, SATTAR J. Enhancing underwater imagery using generative adversarial networks [C]. 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018: 7159-7165.

[3] LEE H S, MOON S W, CHOI S W, et al. Underwater image enhancement using color channel compensation by backscatter light and superpixel dark channel prior[J]. Journal of the Institute of Electronics and Information Engineers, 2020, 57(7): 47-56.

[4] YANG N, ZHONG Q, LI K, et al. A reference-free underwater image quality assessment metric in frequency domain[J]. Signal Processing: Image Communication,

- 2021, 94: 116218.
- [5] 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 ZH P, QU S Y, XING W, et al. An overview of underwater image enhancement methods [J]. CAAI Transactions on Intelligent Systems, 2022, 17 (5): 860-873.
- [6] MA J, FAN X, YANG S X, et al. Contrast limited adaptive histogram equalization-based fusion in YIQ and HSI color spaces for underwater image enhancement[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2018, 32(7): 1854018.
- [7] 王聪, 薛晓军, 李恒, 等. 基于颜色校正和改进二维伽马函数的水下图像增强[J]. 电子测量与仪器学报, 2021, 35(2): 171-178.
WANG C, XUE X J, LI H, et al. Underwater image enhancement based on color correction and improved two-dimensional gamma function [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35 (2): 171-178.
- [8] WONG S L, PARAMESRAN R, TAGUCHI A. Underwater image enhancement by adaptive gray world and differential gray-levels histogram equalization [J]. Advances in Electrical and Computer Engineering, 2018, 18(2): 109-116.
- [9] YUAN J, CAO W, CAI Z, et al. An underwater image vision enhancement algorithm based on contour bougie morphology [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(10): 8117-8128.
- [10] 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. IEEE, 2012: 81-88.
- [11] PENG Y T, CAO K, COSMAN P C. Generalization of the dark channel prior for single image restoration [J]. IEEE Transactions on Image Processing, 2018, 27(6): 2856-2868.
- [12] ANCUTI C O, ANCUTI C, DE VLEESCHOUWER C, et al. Locally adaptive color correction for underwater image dehazing and matching [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017: 1-9.
- [13] ZHU X, LIN M, ZHAO M, et al. Adaptive underwater image enhancement based on color compensation and fusion [J]. Signal, Image and Video Processing, 2023: 1-10.
- [14] SONG W, WANG Y, HUANG D, et al. A rapid scene depth estimation model based on underwater light attenuation prior for underwater image restoration [C]. Advances in Multimedia Information Processing-PCM 2018: 19th Pacific-Rim Conference on Multimedia, Hefei, China, September 21-22, 2018, Proceedings, Part I 19. Springer International Publishing, 2018: 678-688.
- [15] 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.
- [16] WANG Y, ZHANG J, CAO Y, et al. A deep CNN method for underwater image enhancement [C]. 2017 IEEE International Conference on Image Processing (ICIP). IEEE, 2017: 1382-1386.
- [17] 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. 2021, 35 (18): 15853-15854.
- [18] ZHANG W, LIU W, LI L, et al. A framework for the efficient enhancement of non-uniform illumination underwater image using convolution neural network [J]. Computers & Graphics, 2023, 112: 60-71.
- [19] LI J, 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, 2017, 3(1): 387-394.
- [20] HOU L, YU L, TIAN S, et al. FMAGAN: Fusing multiple attention and generative adversarial network to enhance underwater image [J]. Journal of Intelligent & Fuzzy Systems, 2022, 42(3): 2421-2433.
- [21] HAN G, WANG M, ZHU H, et al. UIEGAN: Adversarial learning-based photo-realistic image enhancement for intelligent underwater environment perception [J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, DOI: 10.1109/TGRS.2023.3281741.
- [22] CHEN X, YU J, KONG S, et al. Towards real-time advancement of underwater visual quality with GAN [J]. IEEE Transactions on Industrial Electronics, 2019, 66(12): 9350-9359.
- [23] GUO Y, LI H, ZHUANG P. Underwater image

- enhancement using a multiscale dense generative adversarial network [J]. *IEEE Journal of Oceanic Engineering*, 2019, 45(3): 862-870.
- [24] LI C, ANWAR S, PORIKLI F. Underwater scene prior inspired deep underwater image and video enhancement[J]. *Pattern Recognition*, 2020, 98: 107038.
- [25] WANG Y, GUO J, GAO H, et al. UIEC²-Net: CNN-based underwater image enhancement using two color space [J]. *Signal Processing: Image Communication*, 2021, 96: 116250.
- [26] REN T, XU H, JIANG G, et al. Reinforced swin-convts transformer for simultaneous underwater sensing scene image enhancement and super-resolution [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-16.
- [27] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 10012-10022.
- [28] LIU S, FAN H, LIN S, et al. Adaptive learning attention network for underwater image enhancement[J]. *IEEE Robotics and Automation Letters*, 2022, 7(2): 5326-5333.
- [29] ISLAM M J, XIA Y, SATTAR J. Fast underwater image enhancement for improved visual perception [J]. *IEEE Robotics and Automation Letters*, 2020, 5(2): 3227-3234.
- [30] LU J, LI N, ZHANG S, et al. Multi-scale adversarial network for underwater image restoration [J]. *Optics & Laser Technology*, 2019, 110: 105-113.
- [31] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]. *Proceedings of the IEEE International Conference on Computer Vision*, 2017: 2223-2232.
- [32] HAN J, SHOEIBY M, MALTHUS T, et al. Underwater image restoration via contrastive learning and a real-world dataset [J]. *Remote Sensing*, 2022, 14(17): 4297.
- [33] PARK J, HAN D K, KO H. Adaptive weighted multi-discriminator CycleGAN for underwater image enhancement [J]. *Journal of Marine Science and Engineering*, 2019, 7(7): 200.
- [34] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [C]. *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 3-19.
- [35] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7132-7141.
- [36] QIN Z, ZHANG P, WU F, et al. Fcanet: Frequency channel attention networks [C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 783-792.
- [37] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 11534-11542.
- [38] WANG Z, JI S. Smoothed dilated convolutions for improved dense prediction [C]. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018: 2486-2495.
- [39] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 1125-1134.
- [40] DING X, ZHANG X, MA N, et al. Repvgg: Making vgg-style convnets great again [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 13733-13742.
- [41] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. *Advances in Neural Information Processing Systems*, 2015, 28.
- [42] HAN J, SHOEIBY M, MALTHUS T, et al. Underwater image restoration via contrastive learning and a real-world dataset [J]. *Remote Sensing*, 2022, 14(17): 4297.
- [43] ZHENG C, CHAM T J, CAI J. The spatially-correlative loss for various image translation tasks [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 16407-16417.
- [44] WANG Z, 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.
- [45] PANETTA K, GAO C, AGAIAN S. Human-visual-system-inspired underwater image quality measures [J]. *IEEE Journal of Oceanic Engineering*, 2015, 41(3): 541-551.
- [46] 马红强, 马时平, 许悦雷, 等. 基于深度卷积神经网络的低照度图像增强 [J]. *光学学报*, 2019, 39(2): 99-108.
- MA H Q, MA SH P, XU Y L, et al. Low-light image enhancement based on deep convolutional neural network [J].

Acta Optica Sinica, 2019, 39(2): 99-108.

作者简介



胡雨航, 2021 年于昆明理工大学获得学士学位, 现为昆明理工大学的硕士研究生, 主要研究方向为水下图像处理。

E-mail: huyuhang0219@163.com

Hu Yuhang received his B. Sc. degree from Kunming University of Science and

Technology in 2021. Now he is a M. Sc. candidate at Kunming University of Science and Technology. His main research interest includes underwater image processing.



赵磊(通信作者), 2002 年于昆明理工大学获得学士学位, 2006 年于昆明理工大学获得硕士学位, 现为昆明理工大学实验师, 主要研究方向为图像处理、机器学习。

E-mail: zhaolei@sohu.com

Zhao Lei (Corresponding author)

received his B. Sc. degree from Kunming University of Science and Technology in 2002, M. Sc. degree from Kunming University of Science and Technology in 2006 respectively. Now he is an experimenter in Kunming University of Science and Technology. His main research interests include image processing and machine learning.