

# 控制与决策

Control and Decision

## 动态异构特征融合的水下图像增强算法

韩晓微, 张云泽, 谢英红, 吴宝举, 赵玉莹

引用本文:

韩晓微,张云泽,谢英红,吴宝举,赵玉莹. 动态异构特征融合的水下图像增强算法[J]. *控制与决策*, 2023, 38(6): 1560–1568.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.1826>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于注意力特征融合的无人机多目标跟踪算法

UAV multi-target tracking algorithm based on attention feature fusion

控制与决策. 2023, 38(2): 345–353 <https://doi.org/10.13195/j.kzyjc.2021.1098>

#### 改进残差密集生成对抗网络的红外与可见光图像融合

Infrared and visible image fusion with improved residual dense generative adversarial network

控制与决策. 2023, 38(3): 721–728 <https://doi.org/10.13195/j.kzyjc.2021.1580>

#### 基于三端注意力机制的视网膜血管分割算法

Improved U-Net based on three-terminal attention mechanism for retinal vessel segmentation

控制与决策. 2022, 37(10): 2505–2512 <https://doi.org/10.13195/j.kzyjc.2021.0435>

#### 融合注意力机制的域泛化行人再识别

Domain generalization person re-identification based on attention mechanism

控制与决策. 2022, 37(7): 1721–1728 <https://doi.org/10.13195/j.kzyjc.2020.1844>

#### 基于多尺度残差注意网络的轻量级行人属性识别算法

Lightweight pedestrian attribute recognition algorithm based on multi-scale residual attention network

控制与决策. 2022, 37(10): 2487–2496 <https://doi.org/10.13195/j.kzyjc.2021.0411>

# 动态异构特征融合的水下图像增强算法

韩晓微<sup>1†</sup>, 张云泽<sup>2</sup>, 谢英红<sup>2</sup>, 吴宝举<sup>1</sup>, 赵玉莹<sup>2</sup>

(1. 沈阳大学 科技创新研究院, 沈阳 110044; 2. 沈阳大学 信息工程学院, 沈阳 110044)

**摘要:** 针对水下图像细节模糊和色彩失真严重的问题, 提出一种基于编码解码结构的动态异构特征融合水下图像增强网络. 首先, 设计异构特征融合模块, 将不同级别与不同层次的特征进行融合, 提升网络对细节信息和语义信息的整体感知能力; 然后, 设计新型特征注意力机制, 改进传统通道注意力机制, 并将改进后的通道注意力与像素注意力机制加入异构特征融合过程, 加强网络提取不同浑浊度像素特征的能力; 接着, 设计动态特征增强模块, 自适应扩展感受野以提升网络对图像畸变景物的适应力和模型转换能力, 加强网络对感兴趣区域的学习; 最后, 设计色彩损失函数, 并联合最小化绝对误差损失与结构相似性损失, 在保持图像纹理的基础上纠正色偏. 实验结果表明, 所提出算法可有效提升网络的特征提取能力, 降低水下图像的雾度效应, 提升图像的清晰度和色彩饱和度.

**关键词:** 深度学习; 神经网络; 注意力机制; 异构特征融合; 编码解码结构; 水下图像增强

中图分类号: TP391

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.1826

开放科学(资源服务)标识码(OSID):



引用格式: 韩晓微, 张云泽, 谢英红, 等. 动态异构特征融合的水下图像增强算法[J]. 控制与决策, 2023, 38(6): 1560-1568.

## Underwater image enhancement algorithm based on dynamic heterogeneous feature fusion

HAN Xiao-wei<sup>1†</sup>, ZHANG Yun-ze<sup>2</sup>, XIE Ying-hong<sup>2</sup>, WU Bao-ju<sup>1</sup>, ZHAO Yu-ying<sup>2</sup>

(1. Institute of Innovation Science & Technology, Shenyang University, Shenyang 110044, China; 2. College of Information Engineering, Shenyang University, Shenyang 110044, China)

**Abstract:** Aiming at the problems of blurred details of underwater images and serious color distortion, this paper proposes a dynamic heterogeneous feature fusion underwater image enhancement network based on the autoencoder structure. First, a heterogeneous feature fusion module is designed to integrate different levels and different levels of features to improve the overall perception of detailed information and semantic information of the network. Second, a new feature attention mechanism is designed, the traditional channel attention mechanism is improved, and the improved channel attention and pixel attention mechanism is added to the heterogeneous feature fusion process to strengthen the network's ability to extract pixel features of different turbidity. Then, a dynamic feature enhancement module is designed to adaptively expand the receptive field to improve the network's adaptability to image distortion scenes and model conversion capabilities, and strengthen the network's learning of regions of interest. Finally, the color loss function is designed, and the absolute error loss and the structural similarity loss are jointly minimized, and the color cast is corrected on the basis of maintaining the image texture. The experimental results show that the proposed algorithm can effectively improve the feature extraction ability of the network, reduce the haze effect of underwater images, and improve the clarity and color saturation of the image.

**Keywords:** deep learning; neural networks; attention mechanism; heterogen feature fusion; encoder decoder structure; underwater image enhancement

## 0 引言

复杂的水下环境会导致图像产生色偏、雾度效应、对比度低等问题, 严重阻碍了机器人对水下场景的视觉感知<sup>[1]</sup>. 因此, 有必要对水下图像增强算法进

行研究, 改善水下机器人执行高级视觉任务的性能.

水下图像增强工作可大致分为传统的增强方法和基于深度学习的增强方法. 传统方法又可大致分为基于非物理模型的方法和基于物理模型的图像增

收稿日期: 2021-10-25; 录用日期: 2022-03-15.

基金项目: 国家自然科学基金项目(61873338); 沈阳市科学技术计划项目(21-103-0-15).

责任编辑: 侯忠生.

<sup>†</sup>通讯作者. E-mail: hxw69@163.com.

强方法. 基于非物理模型的方法如直方图均衡化<sup>[2]</sup>、灰度世界假设<sup>[3]</sup>、小波变换<sup>[4]</sup>、自动白平衡<sup>[5]</sup>等通过直接调整图像像素值的方式提升水下图像的质量,但由于没有考虑物理模型,易导致图像颜色偏差大、图像伪影等问题. 基于物理模型的方法包括基于图像去雾模型 Jaffe-McGlamery 的恢复方法<sup>[6-7]</sup>,如暗通道先验<sup>[8-9]</sup>、模糊先验<sup>[10-12]</sup>等. 基于物理模型的方法依赖于假设先验的水下物理成像模型,计算特征的过程非常复杂,且先验的假设较为理想,局限性大,不能很好地实现水下图像色彩校正和移除人造光斑.

近年来,深度学习方法的应用大大改善了低级视觉任务的效果<sup>[13-16]</sup>. 相较于传统方法,基于深度学习的方法优点在于不需要人工提取特征,对图像处理任务的理解更加客观、准确. 深度学习方法可大致分为基于卷积神经网络的方法和基于生成对抗网络的方法<sup>[17-19]</sup>. 基于卷积神经网络的方法使用卷积运算建立端到端的自动特征提取模型,得到水下图像与清晰图像之间的映射关系,进而实现水下图像恢复. 生成对抗网络的方法借鉴博弈论的思想构建生成器和对抗器,二者互相对抗学习,逐步恢复水下图像的质量. 如 Wang 等<sup>[20]</sup>提出了一种基于卷积神经网络的水下图像增强网络 (underwater image enhancement networks, UIE-Net), 该网络由用于颜色校正的子网络 (color correction networks, CC-Net) 和用于提升图像对比度的子网络 (haze removal networks, HR-Net) 构成,分别输出色彩校正的图像和光衰减的透射图,该方法给出了一种像素破坏策略,通过对局部区域进行平滑处理实现水下图像增强,降低了微小纹理和噪声的影响,但恢复后的图像存在细节模糊的问题. Li 等<sup>[21]</sup>提出了一种门控融合卷积神经网络结构,采用多种图像预处理的方式,通过应用白平衡、伽玛校正和直方图均衡将改善后的图像为神经网络提供多输入方式以提升水下图像质量,但恢复后的图像产生伪影,色彩恢复效果差. Sun 等<sup>[22]</sup>提出了一种基于去噪自编码器结构的水下色彩校正网络,编码器逐步提取水下图像特征,解码器通过卷积上采样逐步恢复图像分辨率,从而达到提升水下图像质量的效果,该方法结构简单,未考虑层与层之间的信息交流,退化图像中较为浑浊的区域恢复效果较差. Fabbri 等<sup>[23]</sup>提出了一种基于 Wasserstein GAN 的水下图像增强网络,生成器采用类编码器-解码器结构,判别器基于全卷积网络,该方法使用了大量的卷积,导致网络参数过多,且图像恢复效果不明显. 李庆忠等<sup>[24]</sup>则利用循环对抗网络 CycleGAN,使用对称的 GAN 网络构成一个环形网络,并设计强结构相似性损失函数约束图像纹

理,进行两阶段的训练,基于 GAN 网络的方法不再需要配对水下图像进行训练,但会导致图像失真,且在速度上极为不占优势.

上述方法虽然显著地提高了水下图像质量,但在图像恢复效果上存在边缘丢失、细节模糊、退化严重、区域增强效果差的问题,在网络结构设计和特征提取能力等方面仍有较大的改进空间. 针对上述问题,本文提出一种基于动态异构特征融合的水下图像增强算法. 该算法的主要内容包括:首先,采用异构特征融合的方式将高级与低级特征进行融合,改善信息流动,丰富图像细节信息;同时,改进注意力机制模块,提升网络纠正色偏和去除浑浊度的能力;然后,设计动态特征增强模块,扩大感受野,提升网络对感兴趣区域的特征提取能力;最后,设计损失函数,将色彩损失与最小化绝对误差损失和结构相似性损失相结合,提升模型的拟合能力和增强效果. 实验结果表明,所提出算法能够在提升水下图像清晰度和保证图像纹理结构的同时,恢复更好的色彩效果.

## 1 动态异构特征融合网络

### 1.1 网络结构设计

为解决现有方法存在增强效果不均匀和图像细节丢失问题,本文设计一种端到端的异构特征融合和动态特征增强的水下图像恢复网络. 网络结构如图 1 所示,在编码器-解码器结构的基础上,融入特征注意模块、动态特征增强模块与异构特征融合模块. 在编码器结构中加入改进后的注意力机制从低层到高层逐步提取不同类型的特征;动态特征增强模块在低分辨率空间进行加强特征提取;解码器使用上采样操作将特征向量重构,逐步恢复为相应的清晰图像. 随着网络层深度的加深,浅层特征逐渐退化,使用特征融合模块将低层与高层信息进行融合,改善层之间的信息传递,减少细节特征丢失. 其中,编码器是由 8 倍下采样操作和改进特征注意力机制组成,8 倍下采样由一个步幅为 1 的标准  $7 \times 7$  卷积和 3 个步幅为 2 的标准卷积实现,解码器则使用转置卷积作为相应上采样方法逐步恢复图像分辨率.

#### 1.1.1 特征注意力模块设计

由于大多数编码器-解码器结构在图像处理过程中对整个图像进行无差别对待,导致当图像信息较大时,提取到较多不相关特征,难以适应复杂多变的水下环境,需要添加注意力机制以提高模型的特征提取准确度<sup>[25]</sup>. 为了解决上述问题,本文对通道注意力机制和像素注意力机制进行改进并设计特征注意模块,提升网络对水下图像不同浑浊度区域的适应性. 图 2

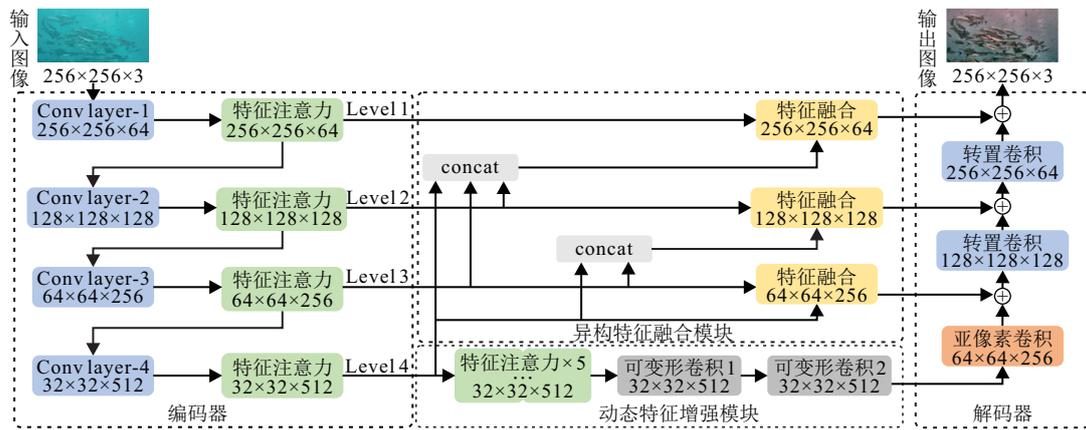


图1 动态异构特征融合下的水下图像增强网络结构

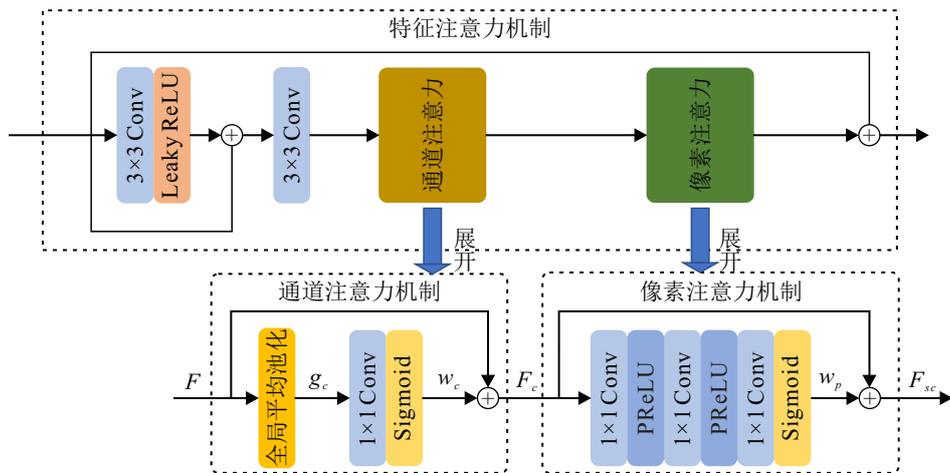


图2 特征注意力模块结构

为特征注意模块结构示意图,设计局部残差结构将上层的输出特征先经过一个卷积,再通过Leaky ReLU激活函数转换为非线性,将输出结果与输入特征叠加,对叠加后的特征再进行一次卷积,并将得到的特征传入改进后的通道注意力机制和像素注意力机制中提取其中更为隐含的映射关系。

### 1)改进通道注意力机制。

以往的通道注意力机制通常对输入特征 $F$ 使用全局平均池化将通道相关的信息转化为通道描述符 $g_c$ ,有

$$g_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F(i, j). \quad (1)$$

在全局平均池化的作用下,特征图的形状从 $C \times H \times W$ 变为 $C \times 1 \times 1$ ,且使用全连接层以获取不同通道的权重,通过降维以降低模型复杂度,将通道特征投影至低维空间后重新映射,此种方法会破坏通道与权重间的直接对应关系,影响通道注意力的预测效果。为提升通道间的相关性,本文将通道注意力模块中的全连接层替换为内核大小为 $k$ 的快速一维卷积,且所有通道共享学习参数,有

$$W_c = \sigma(\text{Conv}_1 D_k(g_c)). \quad (2)$$

其中: $\sigma$ 为Sigmoid激活函数, $\text{Conv}_1 D$ 为1维卷积。设计通道维数 $C$ 与内核大小 $k$ 之间的映射函数如下式所示,自适应地确定内核大小 $k$ :

$$C = \phi(k) = 2^{(\gamma \times k - b)}. \quad (3)$$

给定通道维数 $C$ , $k$ 为

$$k = \Psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}}. \quad (4)$$

其中:将 $\gamma$ 设为2, $b$ 设为1, $\lfloor \cdot \rfloor_{\text{odd}}$ 为 $\zeta$ 的最近奇数。

最后将输入特征 $F$ 与所学习到的通道权重 $W_c$ 逐像素相乘输出通道特征 $F_c$ ,即

$$F_c = F \otimes W_c. \quad (5)$$

### 2)像素注意力机制。

简单的编码器-解码器网络结构对图像内容会进行无差别处理,而不同水下环境所产生的浑浊度并不均匀。因此,作为通道注意力机制的补充,本文设计逐像素注意力机制,使得网络能够关注浑浊度更高的图像区域。

与通道注意力机制类似,将经过通道注意力机

制的特征  $F_c$  作为逐像素注意的输入传入像素注意力层, 像素注意力层为 2 个 PReLU 激活函数和 1 个 Sigmoid 激活函数构成的 3 个卷积层, 有

$$W_p = \sigma(\text{Conv}(\delta(\text{Conv}(\delta(\text{Conv}(F_c)))))). \quad (6)$$

其中: Conv 为  $1 \times 1$  卷积,  $\delta$  为 PReLU 激活函数,  $\sigma$  为 Sigmoid 激活函数. 经过一系列卷积操作, 特征形状由  $C \times H \times W$  变为  $1 \times H \times W$ . 最后, 对输入特征  $F_c$  与所获得的权重  $W_p$  进行逐像素相乘得到输出特征  $F_{sc}$ , 即

$$F_{sc} = F_c \otimes W_p. \quad (7)$$

### 1.1.2 动态特征增强模块设计

由于标准卷积核的采样结构是固定的, 导致其在采样过程中难以适应水下图像中复杂的几何形变. 为降低此种限制, 使得网络特征提取能够适应复杂的水下环境, 本文使用多个特征注意力机制和 2 个可变形卷积层<sup>[26]</sup>设计了动态特征增强模块, 如图 1 所示. 动态特征增强模块可依据水下图像中目标物不同的位置、大小和形状自适应动态调整卷积核的采样位置, 实现更为准确的特征提取. 使用 2 个可变形卷积级联构成动态特征增强模块, 进一步提升采样网络的变换能力.

可变形卷积的特征提取过程增加了偏置区域和偏置向量部分, 标准卷积对输入特征图进行卷积操作后不再直接输出特征图, 而是获得一个与输入特征图大小相同的偏置区域, 然后利用相同的卷积核对新的偏置区域进行特征提取, 得到变形后的特征图.

对于输入的退化图像, 其每个像素点使用  $3 \times 3$  标准卷积计算, 如下式所示:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n). \quad (8)$$

其中:  $p_0$  为输入图像卷积中心的坐标,  $R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}$  定义了卷积感受野的中心点的位置和偏移范围,  $p_n$  列举了  $R$  范围的位置坐标,  $w$  为采样位置的加权和,  $x$  为输入图像特征图,  $y$  为卷积后得到的输出特征映射.

可变形卷积在标准卷积的基础上增加了位置偏置量  $\{V_{p_n} | n=1, 2, \dots, N, N \in |R|\}$ , 其表达式变为

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + V_{p_n}). \quad (9)$$

为尽量避免引入相关性较差的上下文区域特征, 在式 (9) 的基础上添加了权重系数  $\Delta m_n$ ,  $\Delta m_n$  主要用以判断对各采样点位置输入的感兴趣程度, 若对此区域不感兴趣则将其权重系数调整为 0, 其表达式为

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + V_{p_n}) \cdot V_{m_n}. \quad (10)$$

由于偏移量的添加往往会产生浮点类型的数据, 为了能够顺利完成反向传播, 采用双线性插值获得采样点对应的像素, 有

$$x(p) = \sum_q G(q, p) \cdot x(q). \quad (11)$$

其中:  $p$  为添加了偏移后的任意采样位置坐标即式 (10) 中的  $p = p_0 + p_n + V_{p_n}$ ,  $q$  枚举特征映射  $x$  中所有整体空间位置,  $G(\cdot, \cdot)$  为 4 个整数坐标分别对应的权重.

如图 3 所示, 通过使用动态特征增强进行自适应采样, 采样位置不再受限于固定的形状和尺寸, 而是融合了更多的空间结构信息, 更加符合图像中目标物本身的几何形状, 保证图像内容的原有结构, 降低图像恢复过程出现偏差的概率.

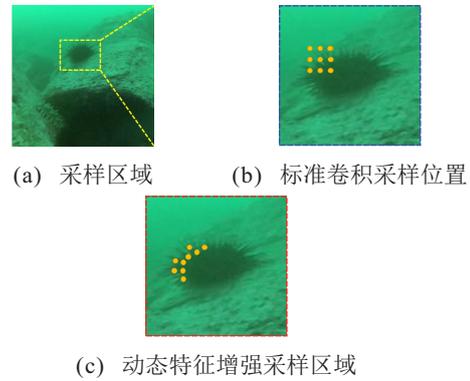


图 3 动态特征增强模块

### 1.1.3 异构特征融合

卷积神经网络可在浅层捕获图像的低级特征, 如边缘和纹理, 然而随着网络层的增加, 浅层特征逐渐退化, 可能会出现低级特征丢失的现象. 以往的融合方式通常使用简单的跳连接将同一级别的特征进行融合, 网络的深层结构与浅层结构之间无法进行信息交流, 易导致信息丢失. 为了解决此问题, 在下采样与上采样之间将高级语义信息嵌入低级特征, 并将低级空间信息嵌入高级特征, 实现异构特征融合. 高级语义嵌入表示为

$$y_i = F(x_i, x_{i+1}, \dots, x_I) + \text{Up}(y_{i+1}). \quad (12)$$

其中:  $y_i$  为第  $i$  级融合特征,  $x_i$  为第  $i$  个通过编码器的特征,  $\text{Up}(\cdot)$  为上采样操作. 目的是将更多的高级特征引入至低级特征. 具体实现如图 4 所示, 首先将高于当前级别的输入特征使用双线性插值调整至统一尺寸后进行堆叠, 然后使用  $3 \times 3$  卷积调整堆叠后的特征通道数与低级别特征一致, 接着使用双线性插值上采样方法将特征图大小调整至与来自下采样的低级特征图相同, 最后将堆叠以及处理后的高级特征图与低级特征图进行逐像素相乘, 实现高级特征语义信息嵌入低级特征.

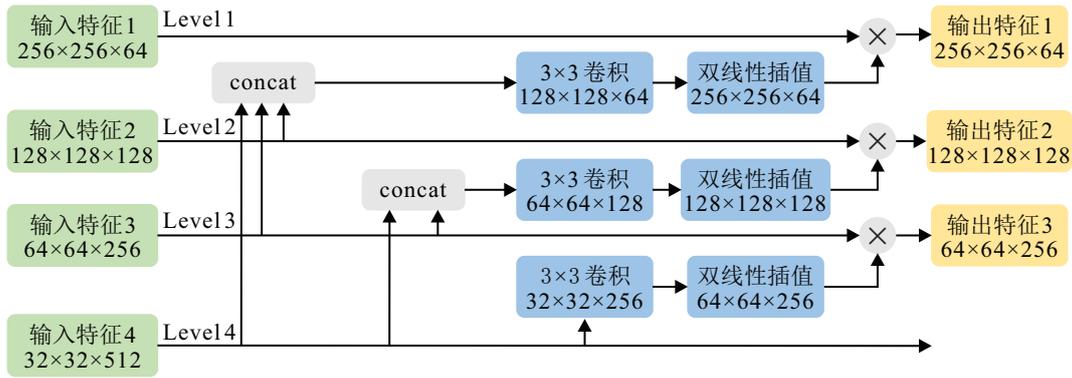


图4 异构特征融合结构设计

另外,由于转置卷积包含特征权重信息,在第1阶段上采样过程中使用无参数上采样方法亚像素卷积<sup>[27]</sup>(Sub-pixel Convolution),通过重塑空间和通道维度的方式扩大特征图,将空间信息引入至高级语义.亚像素卷积表示为

$$I_h = f(I_l) = \varphi(W_l \cdot f_{l-1}(I_l) + b_l). \quad (13)$$

其中: $I_l$ 为输入的低分辨率特征图; $f(\cdot)$ 为特征图由低分辨率到高分辨率的映射函数; $\varphi$ 为周期筛选运算符,将 $H \times W \times C \cdot r^2$ 形状的张量重新排列形状为 $rH \times rW \times C$ 的张量; $W_l$ 和 $b_l$ 分别为权重和偏置.亚像素卷积通过融合 $r^2$ 个与上层特征相同尺寸的特征图,重塑空间向量和通道维数扩大特征图尺寸,配合转置卷积进行上采样可有效地将空间信息引入至高级语义.

## 1.2 损失函数设计

水下图像存在亮度低、图像模糊、颜色偏差等问题,以往的水下图像增强方法仅使用单一的损失函数可能无法达到很好的恢复效果.本文综合考虑最小化绝对误差损失 $L_{l_1}$ 、结构相似性损失 $L_{ssim}$ 和色彩损失 $L_{color}$ ,设计适合水下图像的损失函数.

### 1.2.1 最小化绝对值误差损失

在以往的图像恢复任务中常使用 $L_2$ 损失,但由于其自身缺陷,会过度惩罚较大的误差,降低图像恢复质量.为避免此问题,本文使用最小化绝对误差损失,如下式所示:

$$L_{l_1}(P) = \frac{1}{N} \sum_{p \in P} |x(p) - y(p)|. \quad (14)$$

其中: $p$ 为像素的索引; $P$ 为像素块; $x(p)$ 和 $y(p)$ 分别为增强后图像中的像素块和参考图像中像素的值; $L_{l_1}$ 损失通过逐像素对比求绝对值,保持图像的亮度和颜色不变.

### 1.2.2 结构相似性损失

使用类自编码器结构进行图像重建时,原始图像的结构纹理易发生扭曲.结构相似性<sup>[28]</sup>可通过计算

原始图像与标签图像之间的亮度、对比度和结构3部分的组合差异以衡量图像相似度.在水下图像增强和其他低级视觉任务中,保证图像亮度、对比度和纹理结构的一致性是非常必要的.其中使用均值估计亮度、标准差估计对比度和协方差估计结构相似度,如下式所示:

$$S(p) = \frac{2u_x u_y + C_1}{u_x^2 + u_y^2 + C_1} \cdot \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}. \quad (15)$$

式中: $u_x$ 和 $u_y$ 分别为像素 $x$ 和像素 $y$ 的平均值; $\sigma_{xy}$ 为 $x$ 与 $y$ 的协方差; $\sigma_x^2$ 和 $\sigma_y^2$ 分别为 $x$ 和 $y$ 的方差;为了防止出现分母为0的情况,设置 $C_1 = 0.01$ , $C_2 = 0.02$ .设计结构相似性损失表达式为

$$L_{ssim} = 1 - S(p) = 1 - \frac{2u_x u_y + C_1}{u_x^2 + u_y^2 + C_1} \cdot \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}. \quad (16)$$

### 1.2.3 色彩损失

尽管在 $L_1$ 损失中已经具有了计算颜色误差的功能,但其仅以通过比较像素值不同的方式衡量生成图像与真实情况图像的色彩,不能保证颜色向量具有相同的方向,此种衡量方法仍然可能会导致明显的色差.因此,本文设计了具有颜色角度信息损失函数.通过逐像素计算生成图像与真实图像像素颜色之间的角度并求和的方式对色彩进一步校正.色彩校正损失表达式如下:

$$L_c = \sum_p \angle((\tilde{x}_i)_p, (\tilde{y}_i)_p). \quad (17)$$

其中: $(\cdot)_p$ 为一个像素; $\angle(\cdot)$ 为一个算子,计算两种颜色间的角度,将RGB颜色转为三维向量.式(17)对 $\tilde{x}_i$ 和 $\tilde{y}_i$ 中每个像素正确的颜色向量间的角度求和.

综上,本文设计的总损失函数 $L$ 由3个分量组成,如下式所示:

$$L = \sum_{i=1}^N \omega_1 L_{l_1} + \omega_2 L_s + \omega_3 L_c, \quad (18)$$

其中 $\omega_1$ 、 $\omega_2$ 和 $\omega_3$ 为平衡各损失的权重.

## 2 实验设置

本文所使用的训练集来源于大型水下图像增强数据集 EUVP<sup>[19]</sup>, 从中挑选 6 500 对图像用于网络训练, 其中 6 300 对作为训练集, 200 对作为验证集. 为了客观公正地表明实验结果, 在 EUVP 其余图像中随机挑选 150 对图像作为测试集 Test A, 使用真实水下图像增强基准数据集 UIEB<sup>[21]</sup> 中的 800 对图像作为测试集 Test B. 除上述全参考数据集外, 本文还从无参考真实水下图像数据集 RUIE<sup>[29]</sup> 中挑选 100 张图像作为测试集 Test C.

网络基于 PyTorch 1.2 深度学习框架实现, 使用 NVIDIA GTX 1 080 8 G GPU 进行训练. 为了增加训练数据, 通过随机旋转 90°、180°、270° 和水平翻转以扩充训练数据集. 使用 Adam 优化器, 指数衰减率分别设置为 0.9 和 0.999. 初始学习率设置为 0.000 1, 批量大小设置为 2. 使用余弦退火策略<sup>[30]</sup> 调整学习率. 整个网络在水下图像数据集中训练  $3 \times 10^5$  步, 通过实

验调整损失权重平衡系数  $\omega_1 = 0.2, \omega_2 = 0.8, \omega_3 = 0.2$ , 共使用 9 个特征注意模块.

## 3 实验结果和分析

### 3.1 主观评价

为了直观地显示所提出算法与对比算法在清晰度和色彩校正方面的增强效果, 使用来自 EUVP、UIEB 的全参考测试集 Test A 和 Test B 以及来自 RUIE 的无参考测试集 Test C 对增强算法进行测试. 图 5 和图 6 为不同算法在全参考图像测试集 Test A 与 Test B 数据集上的主观比较, 其中, Test B 中所含图像更接近于真实水下环境; 图 7 为在无参考图像测试集上的比较. 经过对比, 基于非物理模型的 Fusion<sup>[31]</sup> 方法会出现过度增强和过度曝光的问题, 导致图像不真实, 且存在较为严重的虚假视觉感受. 基于物理模型的 UDCP<sup>[8]</sup> 方法不能很好地去掉图像中的雾度效应, 且导致图像变暗, 虽然在一定程度上提升了对比度,

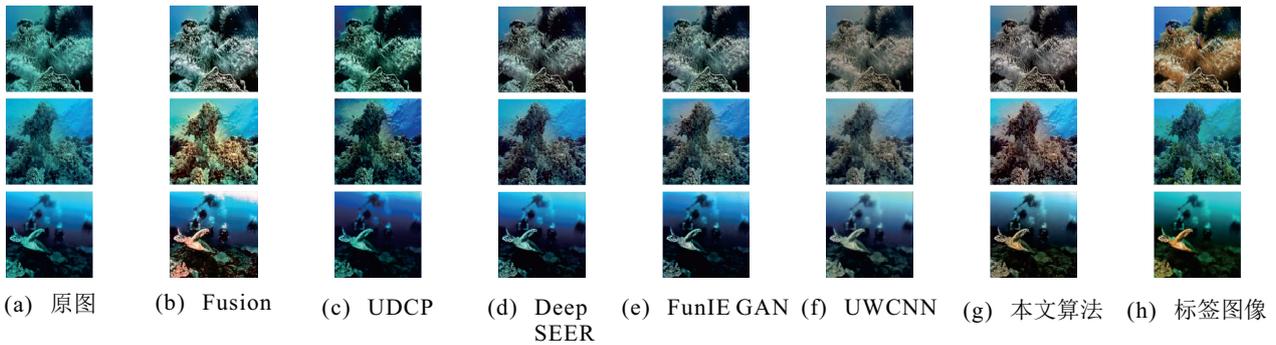


图 5 Test A 增强效果对比

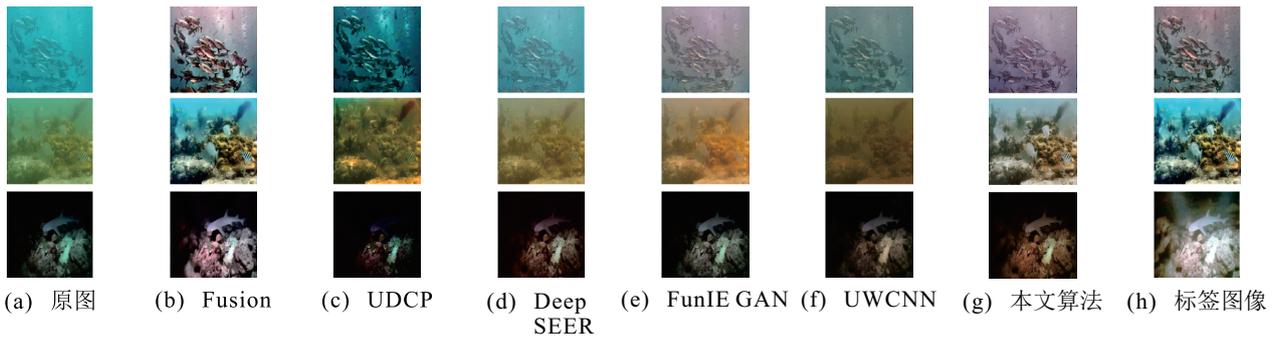


图 6 Test B 增强效果对比

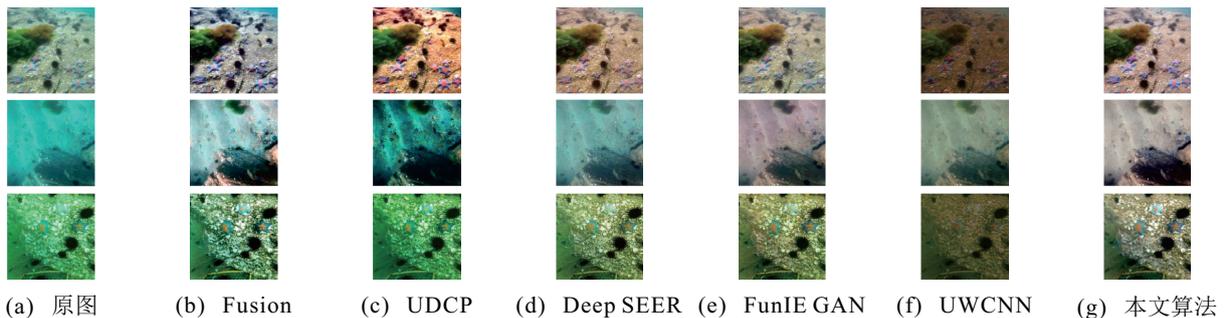


图 7 Test C 增强效果对比

但明显引入了人工颜色产生色彩偏差. 基于深度学习的方法中, Deep SESR<sup>[6]</sup>在一定程度上提升了图像的清晰度, 但效果不明显, 产生了淡黄色的色偏. FunIE-GAN<sup>[19]</sup>利用成对和不成对的数据集训练, 但其网络较浅, 虽然提升了运行速度, 在整体上起到了一定的增强效果, 但局部清晰度不高. UWCNN<sup>[32]</sup>与前文所述的两种基于深度学习的方法存在相同的问题, 虽然在色彩校正方面有一定的提升, 但其图像整体偏暗, 在图像细节增强方面须进一步加强. 与竞争算法相比, 所设计的动态异构特征融合增强网络增强后的图像在全局和局部细节上处理得更加完善, 可更好地去除雾度效果, 使得图像色彩更为均衡, 在视觉效果上也更接近于无水图像.

### 3.2 客观指标

为了更为客观地对算法效果进行分析评价, 除主观感受外还进行了全参考指标评估和无参考指标评估. 对含有标签图像的测试集 Test A 和 Test B, 使用峰值信噪比 (PSNR) 与结构相似性 (SSIM) 进行评估. PSNR 的分数越高表明增强结果在图像内容上与标签图像越接近, 较高的 SSIM 分数表明增强结果与标签图像在结构和纹理上更加接近.

对于不含标签图像的测试集 Test C, 利用无参考水下图像质量指标 UIQM<sup>[33]</sup> 和 UCIQE<sup>[34]</sup> 对增强后的图像进行评估. UIQM 通过度量水下图像色彩差异 (UICM)、水下图像清晰度 (UISM) 和水下图像对比度 (UIConM) 计算加权的方式以评估水下图像退化程度, 其值越大, 表明图像增强效果越好. 如表 1 所示, 所提出算法在测试集 Test A 和 Test B 上均达到最优, 表明所提出算法增强后的图像与标签图像的差异最小, 在提升图像亮度、对比度和保持纹理结构上取得显著效果. 与 Test A 数据集相比, 不同方法于真实水下图像数据集 Test B 的恢复效果相对较差, 相较于对比方法, 所提出算法在真实水下图像上的恢复效果仍然有明显提升. 另外, 所提出算法在测试集 Test C 上的 UIQM 指标稍弱于 FunIE-GAN 方法, UCIQE 达到了最优, 表明所提出方法在提升图像清晰度、色彩饱

表 1 不同方法全参考指标对比

方法	Test A		Test B		Test C	
	PSNR	SSIM	PSNR	SSIM	UIQM	UCIQE
Fusion	17.9	0.73	17.41	0.87	4.71	0.77
UDCP	16.28	0.70	13.65	0.46	2.13	0.69
DEEP	22.35	0.83	18.16	0.74	6.65	0.80
FunIE-GAN	25.66	0.88	17.95	0.75	6.82	0.62
UWCNN	23.62	0.82	14.82	0.61	6.36	0.57
本文方法	27.58	0.94	20.58	0.78	6.73	0.88

和度等方面取得了较好的恢复效果, 更加符合人类视觉感受.

### 3.3 消融实验

为了表明所提出网络各组件的有效性, 在保证其他实验条件相同的情况下, 分别对特征注意模块 CP (channel pixel)、动态特征增强模块 DFE (dynamic feature enhancement)、异构特征融合模块 HFF (heterogeneous feature fusion) 与损失函数部分进行消融实验. 由编码器、解码器、跳连接构成基础网络 (baseline), 将不同模块逐步添加至基础网络. 在不同模块的消融实验中, 仅使用  $L_1$  损失对网络进行训练, 使用 EUVP Image 数据集进行训练和测试, 实验步数设置为  $2 \times 10^5$ . 在仅使用  $L_1$  损失的基础上逐步添加结构相似性损失和色彩损失对网络进行训练, 验证所设计的联合损失函数的有效性.

#### 3.3.1 网络各组件的消融研究

表 2 为模型在测试集 Test A 上的定量对比结果, 由表 2 中 PSNR 和 SSIM 指标可见, 添加所设计模块后可有效地提升基础模型的性能.

表 2 Test A 消融实验参考指标对比

索引	CP	DEF	HFF	PSNR	SSIM
0				25.29	0.83
1	✓			25.74	0.83
2		✓		25.63	0.85
3	✓	✓		25.85	0.89
4	✓	✓	✓	26.63	0.92

为了直观地显示特征注意模块的有效性, 将测试图像输入含有特征注意模块网络结构训练好的模型, 取出图 1 中 Level 4 阶段经特征注意模块后的权重数据生成特征权重图. 如图 8 所示, 通过红色区域可见, 网络能够较多地关注图像中更为浑浊的区域.

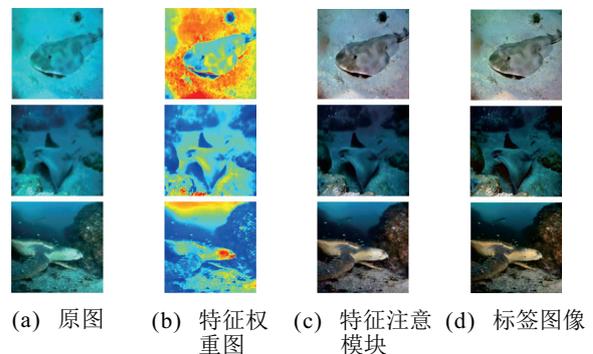


图 8 注意力机制特征权重

图 9 为有无动态特征模块网络增强的效果对比, 可以看出添加动态特征增强模块有利于保持图像边缘结构和纹理, 且细节恢复更加准确、生动.

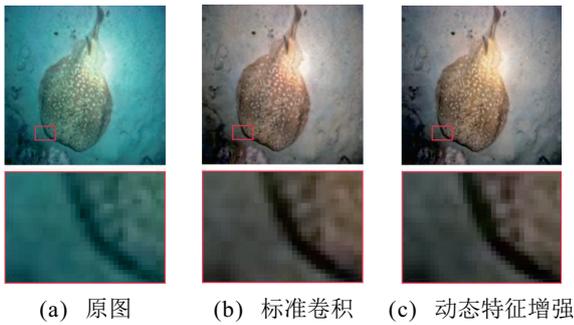


图9 有无动态特征增强效果对比

### 3.3.2 损失函数的消融研究

为了表明所使用多项损失函数的有效性,所设计网络结构逐步对第1.2节所提出损失函数进行消融实验. 经实验测试,图10为在网络训练过程中不同损失函数的变化曲线. 由图10可见,与仅使用 $L_1$ 损失相比,使用多项损失函数可以有效地降低网络的收敛速度和图像差异.

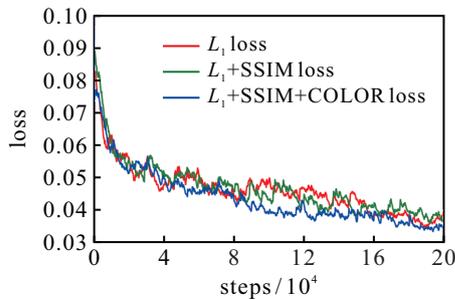


图10 不同损失对比曲线

由表3可见,添加色彩损失后,PSNR和SSIM指标均有提升,表明所设计色彩损失函数可有效地提升网络恢复水下图像色偏的能力.

表3 不同损失参考指标对比

$L_1$	SSIM	COLOR	PSNR	SSIM
✓			26.44	0.92
✓	✓		27.34	0.93
✓	✓	✓	27.59	0.93

## 4 结论

本文提出了一种异构特征融合和动态特征增强的水下图像增强网络. 该网络对编码器-解码器结构进行改进,设计并添加了特征注意模块、动态特征增强模块和异构特征融合模块. 其中:特征注意模块用于辅助编码器和动态特征增强模块加强网络对图像中不同浑浊度区域像素的恢复能力;动态特征增强模块可扩展卷积感受野,保证图像的边缘和纹理结构不发生改变;采用异构特征融合的方式融合网络提取到的低级特征和高级特征,改善各模块间的信息交流,保留更多图像细节信息. 最后,设计一种具有色彩损失的多项损失函数,提升网络的色彩校正能力. 实

验结果表明,所提出算法可有效地提升图像清晰度,校正色彩失真,保证图像细节和纹理,实现水下图像增强.

### 参考文献(References)

- [1] 郭戈, 王兴凯, 徐慧朴. 基于声呐图像的水下目标检测、识别与跟踪研究综述[J]. 控制与决策, 2018, 33(5): 906-922. (Guo G, Wang X K, Xu H P. Review on underwater target detection, recognition and tracking based on sonar image[J]. Control and Decision, 2018, 33(5): 906-922.)
- [2] Reza A M. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement[J]. The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology, 2004, 38(1): 35-44.
- [3] Buchsbaum G. A spatial processor model for object colour perception[J]. Journal of the Franklin Institute, 1980, 310(1): 1-26.
- [4] Singh G, Jaggi N, Vasamsetti S, et al. Underwater image/video enhancement using wavelet based color correction (WBCC) method[C]. IEEE Underwater Technology. Chennai, 2015: 1-5.
- [5] Liu Y C, Chan W H, Chen Y Q. Automatic white balance for digital still camera[J]. IEEE Transactions on Consumer Electronics, 1995, 41(3): 460-466.
- [6] 姚钦舟, 庄苏锋, 屠大维, 等. 水下透视投影图像非线性畸变校正方法[J]. 仪器仪表学报, 2020, 41(2): 176-183. (Yao Q Z, Zhuang S F, Tu D W, et al. Non-linear distortion correction method for underwater perspective projection image[J]. Chinese Journal of Scientific Instrument, 2020, 41(2): 176-183.)
- [7] 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 S, Chi K C, Tang Y D. Underwater image sharpening based on fusion of restored structure and enhanced texture[J]. Control and Decision, 2022, 37(3): 635-644.)
- [8] He K M, Sun J, Tang X O. Single image haze removal using dark channel prior[C]. IEEE Transactions on Pattern Analysis and Machine Intelligence. Piscataway: IEEE, 2011, 33(12): 2341-2353.
- [9] 安冬, 国凌明, 邵萌, 等. 基于暗通道先验的自适应超像素去雾算法[J]. 控制与决策, 2020, 35(8): 1929-1934. (An D, Guo L M, Shao M, et al. An adaptive superpixel defogging algorithm based on dark channel prior[J]. Control and Decision, 2020, 35(8): 1929-1934.)
- [10] Peng Y T, Zhao X Y, Cosman P C. Single underwater image enhancement using depth estimation based on blurriness[C]. IEEE International Conference on Image Processing. Quebec City, 2015: 4952-4956.
- [11] 杨淼, 纪志成. 基于模糊形态筛和四元数的水下彩色图像增强[J]. 仪器仪表学报, 2012, 33(7): 1601-1605.

- (Yang M, Ji Z C. Underwater color image enhancement based on quaternion and fuzzy morphological sieves[J]. Chinese Journal of Scientific Instrument, 2012, 33(7): 1601-1605.)
- [12] Dong C, Loy C C, He K M, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [13] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]. IEEE International Conference on Computer Vision. Venice, 2017: 2242-2251.
- [14] Ren W Q, Liu S F, Ma L, et al. Low-light image enhancement via a deep hybrid network[J]. IEEE Transactions on Image Processing, 2019, 28(9): 4364-4375.
- [15] Guo C L, Li C Y, Guo J C, et al. Zero-reference deep curve estimation for low-light image enhancement[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2020: 1777-1786.
- [16] Islam M J, Luo P G, Sattar J. Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception[J/OL]. 2020, arXiv: 2002.01155.
- [17] 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, 2018, 3(1): 387-394.
- [18] Lu J Y, Li N, Zhang S Y, et al. Multi-scale adversarial network for underwater image restoration[J]. Optics & Laser Technology, 2019, 110: 105-113.
- [19] 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.
- [20] Wang Y, Zhang J, Cao Y, et al. A deep CNN method for underwater image enhancement[C]. IEEE International Conference on Image Processing. Beijing, 2017: 1382-1386.
- [21] Li C, Guo C, Ren W, et al. An underwater image enhancement benchmark dataset and beyond[J]. IEEE Transactions on Image Processing, 2019, 29: 4376-4389.
- [22] Sun X, Liu L P, Dong J Y. Underwater image enhancement with encoding-decoding deep CNN networks[C]. IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation. San Francisco, 2017: 1-6.
- [23] Fabbri C, Islam M J, Sattar J. Enhancing underwater imagery using generative adversarial networks[C]. IEEE International Conference on Robotics and Automation. Brisbane, 2018: 7159-7165.
- [24] 李庆忠, 白文秀, 牛炯. 基于改进CycleGAN的水下图像颜色校正与增强[J]. 自动化学报, 2020, DOI: 10.16383/j.aas.c200510.
- (Li Q Z, Bai W X, Niu J. Underwater image color correction and enhancement based on improved cycle-consistent generative adversarial networks[J]. Acta Automatica Sinica, 2020, DOI: 10.16383/j.aas.c200510.)
- [25] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [26] Zhu X Z, Hu H, Lin S, et al. Deformable ConvNets V2: More deformable, better results[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2019: 9300-9308.
- [27] Shi W Z, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]. IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 1874-1883.
- [28] Zhao H, Gallo O, Frosio I, et al. Loss functions for image restoration with neural networks[J]. IEEE Transactions on Computational Imaging, 2017, 3(1): 47-57.
- [29] Liu R S, Fan X, Zhu M, et al. Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(12): 4861-4875.
- [30] He T, Zhang Z, Zhang H, et al. Bag of tricks for image classification with convolutional neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2019: 558-567.
- [31] Zhang W H, Li G, Ying Z Q. A new underwater image enhancing method via color correction and illumination adjustment[C]. IEEE Visual Communications and Image Processing. Petersburg, 2017: 1-4.
- [32] Li C Y, Anwar S, Porikli F. Underwater scene prior inspired deep underwater image and video enhancement[J]. Pattern Recognition, 2020, 98: 107038.
- [33] Panetta K, Gao C, Agaian S. Human-visual-system-inspired underwater image quality measures[J]. IEEE Journal of Oceanic Engineering, 2016, 41(3): 541-551.
- [34] Yang M, Sowmya A. An underwater color image quality evaluation metric[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2015, 24(12): 6062-6071.

## 作者简介

韩晓微(1969—),男,教授,博士,从事机器人、计算机视觉等研究, E-mail: hxw69@163.com;

张云泽(1996—),男,硕士生,从事深度学习、计算机视觉的研究, E-mail: zhyz96@163.com;

谢英红(1976—),女,教授,博士,从事模式识别、计算机视觉等研究, E-mail: xieyinghong@163.com;

吴宝举(1980—),男,副教授,博士,从事水下机器人等研究, E-mail: 68366487@qq.com;

赵玉莹(1998—),女,硕士生,从事深度学习、计算机视觉的研究, E-mail: 2872585913@qq.com.