

文章编号 1004-924X(2024)13-2112-16

纹理感知联合颜色直方图特征的水下图像增强

袁国铭*, 刘海军, 李晓丽, 张瑞蕾, 单维锋
(防灾科技学院 应急管理学院, 河北 三河 065201)

摘要: 基于深度学习的水下图像增强方法在视觉效果上表现出色,但已有的端到端网络较少针对水下场景中常见的颜色失真和纹理模糊来精心构建网络结构。为了提高网络的性能,提出了纹理感知联合颜色直方图特征的水下图像增强网络,它包含纹理感知网络,颜色直方图特征提取网络以及颜色纹理融合的水下图像增强网络。纹理感知网络的构建引入变形 Transformer 模块,该模块利用空间感知的变形卷积设计 Transformer 模块的多头注意力机制,有效地将变形卷积的几何感知能力与 Transformer 模块的全局语义捕获能力相结合来提取纹理特征。颜色直方图特征提取网络认为直方图中包含丰富的颜色信息,利用真实水下图像的直方图来监督该网络提取颜色特征。最后,利用所提的颜色纹理融合模块将前两个网络提取的纹理和颜色特征融合,并注入到后续的颜色纹理融合的水下图像增强网络中,以实施水下图像增强,该操作不但有效地保留了纹理结构,校正了失真的颜色,还使得纹理和颜色信息保存一致。实验结果表明,本文算法较已有的水下图像增强算法,具有更好的视觉增强效果,水下图像质量测量指标(Underwater Image Quality Metric, UIQM)提高 10%,单张图像的运行时间缩短了 9%,仅为 0.051 s,满足工程实践中水下视觉增强任务要求。

关键词: 图像增强;多头注意力机制;纹理感知;卷积神经网络;深度学习

中图分类号: TP394.1; TH691.9 **文献标识码:** A **doi:** 10.37188/OPE.20243213.2112

Underwater image enhancement using joint texture perception and color histogram features

YUAN Guoming*, LIU Haijun, LI Xiaoli, ZHANG Ruilei, SHAN Weifeng

(Department of emergency management, Institute of Disaster Prevention, Sanhe 065201, China)

* Corresponding author, E-mail: scdyuan@126.com

Abstract: While deep learning methods show promising visual results, current end-to-end networks often lack tailored architectures to address common issues like color distortion and texture blurriness. To improve their effectiveness, we propose an underwater image enhancement network that utilizes joint texture perception and color histogram features. The network comprises a texture-aware module, a color histogram extraction module, and a color-texture fusion enhancement module. The texture-aware network incorporates a deformable transformer module, leveraging spatially aware deformable convolution to enhance multi-head attention and extract texture features. The color histogram extraction module harnesses histograms from real underwater images to compute the loss function. Subsequently, the color-texture fusion module merges the color and texture features, which are then processed by the enhancement network to produce the final results. This approach effectively preserves texture structures, corrects color distortion.

收稿日期: 2023-11-01; 修订日期: 2023-12-12.

基金项目: 中央高校基金资助项目(No. ZY20180223); 河北省自然科学基金资助项目(No. D202251200)

tions, and maintains information consistency. Extensive experiments demonstrate that our method surpasses existing underwater image enhancement algorithms, achieving a 10% increase in the UIQM metric and reducing processing time to just 0.051 s per image. Our model successfully meets the demands of underwater visual enhancement tasks.

Key words: image enhancement; multi-head attention mechanism; texture perception; convolutional neural network(CNN); deep learning

1 引言

随着机器视觉技术的发展,水下机器人广泛用于海洋资源的开发和利用。水下图像作为水下机器人的视觉输入,其质量会影响机器人对环境的感知及后续水下作业的准确度。然而,受水下极端环境,如光的衰减,折射及吸收等的限制,捕获的水下图像往往具有清晰度差,色彩失真,细节模糊等退化问题,严重影响了机器人的水下视觉质量。因此,对捕获的水下图像进行图像增强,改善图像质量对于海洋勘探具有重要意义^[1]。

水下图像增强的目标是将退化的水下图像增强为清晰图像,可分为物理模型增强法,非物理模型增强法和深度学习增强法。物理模型法将图像增强看作水下物理成像模型的逆向求解过程,通过估计模型的参数,逆向推导模型中未退化的清晰图像。经典的水下暗通道先验(Underwater Dark Channel Prior, UDCP)利用水下图像蓝绿通道的局部邻域里,总有像素灰度值趋于零的先验,来估计物理模型中的透射率^[2]。水下光衰减先验(Underwater Light Attenuation Prior, ULAP)利用场景深度与蓝绿通道最大灰度值差的关系来估计物理模型中的景深信息^[3]。然而,受水下场景复杂性的影响,基于先验的物理模型变量估计易失准,算法稳定性差。

非物理模型增强方法不考虑水下图像的物理成像过程,而是通过修改像素的灰度值来增强视觉效果。Li等提出了直方图拉伸及白平衡融合的水下图像增强方法(Hybrid Framework for Underwater Image Enhancement, HUIE)^[4],它通过提高图像的灰度动态范围,达到增强对比度的效果,但处理后的图像易丢失细节。胡等提出了颜色平衡和多尺度融合的水下图像增强方法,它首先利用颜色平衡方法校正图像颜色,再利用自

适应直方图均衡化方法增强 L 通道的对比度,最后将颜色校正的图像及对比度增强后的图像按权重图进行多尺度融合^[5]。非物理模型法具有简洁高效的优势,但水下环境复杂多变,仅靠修改像素的灰度值难以从退化的水下图像中恢复出高质量的清晰图像。

近年来,深度学习算法广泛用于水下图像增强并取得了较好的效果,如生成对抗网络(Generative Adversarial Network, GAN)及卷积神经网络(Convolutional Neural Network, CNN),它们通过一步式端到端的网络结构来实现水下图像增强,能有效避免物理增强模型中复杂的参数估计步骤。为提高网络对不同降质图像的适应性和稳定性,Anwar等通过合成10种Jerlov水类型图像的训练集,来训练编解码网络(Underwater Enhancement CNN, UWCNN),以达到有效增强不同水类型图像效果^[6]。Li等创建了包含光照不均,多样水类型,纹理模糊等不同降质程度的真实水下图像训练集,来训练水下图像增强网络(Underwater Image Enhancement Network, WaterNet)^[7]。Wang等提出了水下生成对抗网络(UnderWater Generative Adversarial Network, UWGAN),它利用陆地图像及其深度图生成与真实水下世界接近的图像,并用该图像训练U形网络来实现图像增强^[8]。尽管丰富训练数据集的多样性可以有效地提高网络的稳定性,但这种端到端的网络并未专门针对颜色失真和纹理模糊的水下图像降质特性来设计网络结构,降低了网络性能。为此,Ma等提出了基于小波的双路径水下图像增强网络(Wavelet-based Dual-stream Network, WDN),它利用编解码器分别对小波分解的高频率细节和低频颜色信息分别进行处理^[9]。Liu等提出自适应学习注意网络(Adaptive Learning Attention Network, ALAN)^[10],它利用注意模块关注照明及颜色信息,利用自适应学

习模块学习纹理特征,取得了较好的增强效果。受此启发,本文进一步针对颜色失真和纹理模糊等降质问题来设计网络结构。

近年来,Transformer在低级别视觉任务中表现出极强的特征提取能力,通过Transformer机制捕获特征间的长范围依赖关系,提高模型对图像细节的重建能力。Ren等提出了Swin-Conv Transformer水下图像增强网络,它通过结合局部注意力机制和全局注意力机制的方式来锐化纹理^[11]。Peng等进一步提出了U形Transformer(U-T)的水下图像增强网络,它结合多尺度特征融合模块和全局特征构建模块来加强网络对颜色和空间区域的关注^[12]。尽管,Transformer在水下图像增强任务中取得巨大进展,但仍无法有效感知水下场景的纹理细节,校正颜色,实施颜色纹理信息一致的图像增强。

本文提出了纹理感知联合颜色直方图特征的水下图像增强网络。它包含纹理感知网络,颜色直方图特征提取网络以及颜色纹理融合的水下图像增强网络。纹理感知网络所提的变形Transformer模块构建而成,不同于传统的Transformer模块利用线性变换来计算多头注意力机制中的查询(Q)、键(K)和数值(V),本文用空间感知的变形卷积作为Transformer模块多头注意力机制中Q,K,V的提取器,有效地将变形卷积的几何感知能力与Transformer模块的全局语义捕获能力相结合来感知并提取纹理特征,解决已有Transformer模块无法感知纹理的问题。颜色直方图特征提取网络认为直方图中包含丰富的颜色信息,它利用真实水下图像的直方图来计算

该网络的损失函数,以便提取逼近真实图像的颜色特征。最后,将前两个网络提取的纹理和颜色特征进一步融合用于水下图像增强。不同于已有的融合策略将颜色和纹理特征串联或相加,提出的颜色纹理融合模块通过计算颜色和纹理的关系矩阵,融合纹理和颜色特征,并将融合特征用于后续颜色纹理融合的水下图像增强网络实施增强。该操作不但能够保留纹理,校正颜色,还能在确保纹理和颜色信息一致的情况下有效实施增强。实验证明,本文算法较已有的水下图像增强算法具有更强视觉增强效果,能有效用于水下机器视觉任务。

2 原 理

纹理感知联合颜色直方图特征的水下图像增强网络的结构如图1所示,包括颜色直方图特征提取网络以及颜色纹理融合的水下图像增强网络。纹理感知网络在变形Transformer模块的帮助下,能有效感知纹理,从水下图像的灰度图中提取纹理特征 V_1, V_2, V_3 。颜色直方图特征提取网络用真实水下图像的直方图计算损失函数,并在其监督下从高斯模糊的水下图像中提取颜色特征 C_1, C_2, C_3 。颜色纹理融合水下图像增强网络以纹理感知网络的输出和原始水下图像作为输入。使用所提的颜色纹理融合模块将纹理信息 V_1, V_2, V_3 和颜色特征 C_1, C_2, C_3 融合为 E_1, E_2, E_3 ,并注入到颜色纹理融合的水下图像增强网络中,用于水下图像增强。为了增强网络的可解释性,将可视化纹理,颜色及其融合特征。

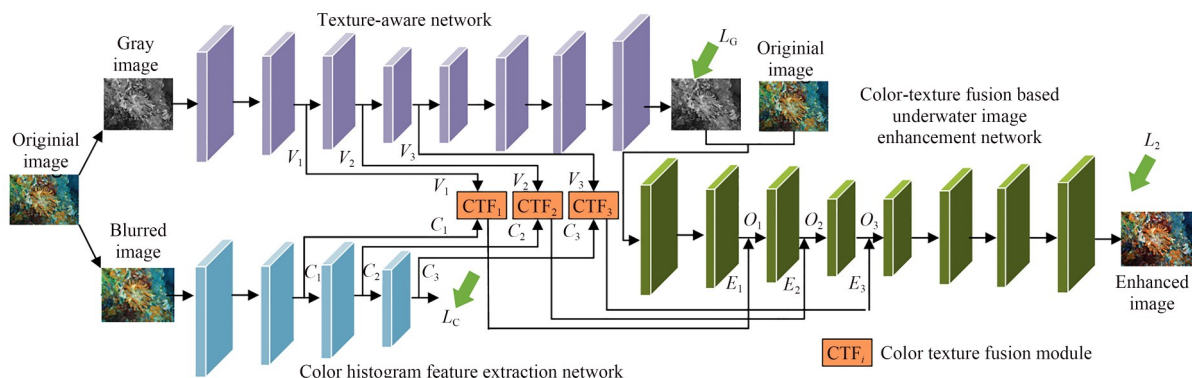


图1 纹理感知联合颜色直方图特征的水下图像增强网络结构

Fig. 1 Architecture of underwater image enhancement using joint texture perception and color histogram features

CAM(Class Activation Mapping)是一种可视化卷积神经网络注意力机制的方法,它能帮助理解模型在分类任务中所关注的区域。通过线性融合网络层输出的各通道特征,并实施归一化及插值处理,CAM可以得到一张和输入图尺寸相等的热力图,该图的像素取值为0~1对应网络对区域的响应强度,值越高响应强度越大。为了更直观地表示,一般将灰度CAM转化为伪彩色的CAM,且颜色越红响应强度越大,颜色越深其响应强度越小。考虑到纹理感知网络,颜色直方图特征提取网络,及颜色纹理融合水下图像增强网

络分别关注纹理,颜色和融合特征,因此可使用CAM来实现特征可视化。如图2所示,图2(a)为UIEB中的一张真实水下图像,图2(b)为高斯模糊结果,图2(c)为UIEB的增强结果。

由图2(d)可见,纹理感知网络所提取的纹理特征 V_3 重点关注图2(a)中珊瑚和鱼等纹理细节丰富的区域,而颜色直方图特征提取网络提取的颜色特征 C_3 (见图2(e))重点关注图2(a)中颜色亮度较大的区域如背景中的鱼以及需要颜色校正的区域如珊瑚,融合后的结果 E_3 (见图2(f))则同时关注了这两类的区域。

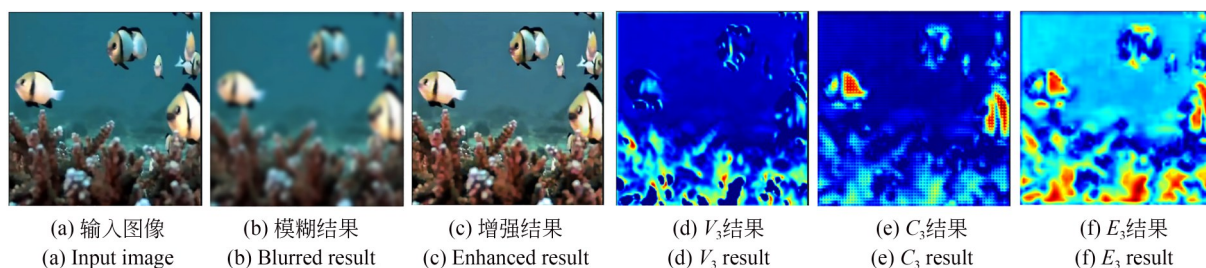


图2 V_3 , C_3 和 E_3 特征的可视化结果

Fig. 2 Visual results of V_3 , C_3 and E_3 features

2.1 纹理感知网络

对于给定的水下图像,纹理结构提取网络的目标是提取不包含颜色的纹理和结构信息。已有方法利用编码解码网络从灰度图像中提取纹理特征^[10],然而网络中所用的卷积层对全局语义关系的捕获能力和物体的几何信息感知能力均较弱,纹理信息提取的准确度不高。为此,本文在编码解码框架中引入可变形Transformer模块(Deformable Transformer Block, DTB)。DTB模块的灵感来源于已有的transformer模块和可变形卷积模块(Deformable Convolution, DC)。相比于传统的卷积层,2017年Vaswani提出的transformer模块得益于能捕获特征间长范围依赖关系的多头注意力机制,具有较强的全局语义关系捕获能力,在自然语言处理乃至机器视觉任务上均表现出卓越的性能。另一方面,可变形卷积由于能有效适应的物体几何变化,具有较强的几何信息感知能力。因此,DTB用空间感知的可变形卷积作为transformer多头注意力机制中查询(Q)、键(K)和数值(V)的特征提取器,能将DC的几何感知力和Transformer的全局语义关

系捕获力相结合,为网络提供丰富的空间上下文信息,达到有效提取纹理特征的目的。

2.1.1 可变形Transformer模块

2.1.1.1 传统Transformer模块结构

Vaswani于2017年首次提出Transformer框架,它在机器翻译、文本生产,乃至机器视觉领域都取得了极大的成就。机器视觉任务中,传统的Transformer模块是由多个Transformer block堆叠而成。单个Transformer block结构如图3(a)所示,它包含多头注意力机制,层归一化和前馈神经网络。其中,前馈神经网络较为简单,可使用经典的线性变换,GELU激活函数及残差连接组成,其结构如图3(b)所示。

多头自注意力机制(Multi-head Attention, MA)是Transformer block的重点。众所周知,MA由多个独立的自注意力子层(Self Attention, SA)构成,如图3(a)所示。每个自注意力子层用于计算输入序列中每个元素与其他元素之间相关性权重,从而通过权重量化表示该元素与其他元素间的关联程度,计算出像素间长范围依赖关系。如图3(b)所示,在具体实现时,每个自注意

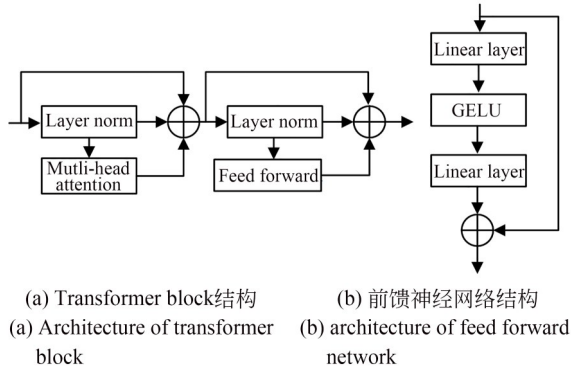


图3 Transformer模块

Fig. 3 Transformer module

力头都有自己的查询 Queries(Q), 键 Key(K) 和值 Values(V) 向量, 用于计算相关性权重。输入特征 F_{in} 在进入卷积层后, 由三个线性层映射为 Q, K, V :

$$Q = F_{in} \cdot W^Q, K = F_{in} \cdot W^K, V = F_{in} \cdot W^V, \quad (1)$$

式中 W^Q, W^K, W^V 为线性变换的权重矩阵。通过查询 Q, K 之间的匹配度计算权重, 并用于加权求和 V , 得到输出的注意力值:

$$F_{out} = \text{Attention}(Q, K, V) = \delta(QK^T)V, \quad (2)$$

式中 δ 代表 softmax 函数。

如图 4(a) 所示, 将多个独立的自注意力子层组合在一起构成多头自注意力机制, 可增强模型的表达能力, 即将 F_{in} 输入到每个自注意力子层, 并将每个自注意力子层输出的向量拼接, 实施线性变换, 得到最终多头自注意力机制的输出。由于每个自注意力层可捕捉输入序列元素间的某种语义关系和依赖性, 多头自注意力机制就可捕

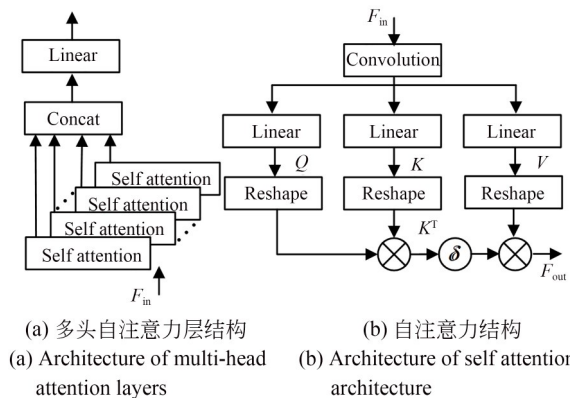


图4 多头自注意力机制

Fig. 4 Multiple-head self attention mechanism

捉多种不同关系和依赖性, 在自然语言处理和计算机视觉等任务中非常有用。然而, 传统的 Transformer block 的多头注意力机制虽能捕获全局语义, 但却不能感知几何信息的模块, 无法捕获上下文中的纹理细节。因此, 本文提出可变形 DTB 模块来解决该问题。

2.1.1.2 空间感知可变形卷积的多头注意力机制

DTB 定义为基于空间感知可变形卷积 (Space Aware Deformable Convolution, SADC) 构建的 Transformer block, 即利用空间感知可变形卷积来计算 Transformer block 多头注意力机制的 Q, K 和 V 。总体来说, DTB 的框架结构与传统 Transformer 一致 (如图 2), 其多头注意力机制中自注意力层的结构如图 5 所示。传统 Transformer block 中 Q, K 和 V 的值由线性变化计算得到 (见式 (1) 和图 3(b)), 与之不同, DTB 用能感知物体几何变化的 SADC 来感知上下文纹理信息, 计算 Q, K 和 V 值。

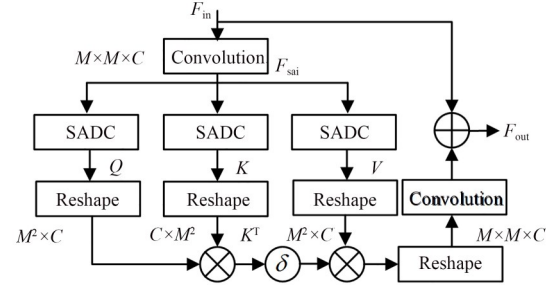


图5 可变形 Transformer block 多头注意力机制中自注意力层结构 (其中, $M \times M$ 为特征的长 $\times$ 宽, C 为通道数, Reshape 为形状重塑层)

Fig. 5 Architecture of self attention in multi-head attention mechanism of deformable transformer block (where $M \times M$ represents the length $\times$ width, C is the channel number, reshape denotes the reshape layer)

假设输入多头注意力机制的特征为 F_{in} , 则单个自注意力层的 Q, K 和 V 计算如下:

$$F_{sai} = f_{conv}(F_{in}), \quad (3)$$

$$Q = f_{sa}(F_{sai}), K = f_{sa}(F_{sai}), V = f_{sa}(F_{sai}), \quad (4)$$

式中: f_{conv} 为核大小为 1×1 的卷积层函数, f_{sa} 为 SADC 函数。由此可见, F_{in} 在进入卷积层后得到特征 F_{sai} , 而 F_{sai} 则分别进入 3 个 f_{sa} , 得到 DTB 中

的 Q, K 和 V 值。由此,自注意力层的输出由下式得到:

$$F_{\text{out}} = F_{\text{in}} + f_{\text{conv}}(\delta(QK^T) \cdot V), \quad (5)$$

式中: f_{conv} 和 F_{in} 的定义与式(3)保持一致, δ 代表 softmax 函数层, F_{out} 代表自注意力层的输出。实验证明,使用 SADC 计算 Q, K 和 V 值,能将 DC 的几何感知力和 Transformer 的全局语义关系捕获力相结合,为网络提供丰富的空间上下文纹理信息。

2.1.1.3 空间感知可变形卷积

由上述分析可知, SADC 是 DTB 感知几何信息的关键。实际上,传统的可变形卷积(Deformable Convolution, DC)已被证明能够适应输入特征中的几何变化,但其偏移量和调制值可能超出相关区域,从而干扰特征的传播和模型的性能。为了使 DC 的注意力集中在空间相关的区域上,提出了空间感知的可变形卷积 SADC,它利用空间注意力机制模块提供与纹理结构相关的偏移量和调制值,使 DC 具有更好的几何信息感知能力。

SADC 的结构如图 6 所示,输入特征在经过一个卷积层后得到特征 F_p , F_p 进一步进入空间注意力机制模块。该模块利用最大池化层和平均池化层计算空间注意力,计算过程如下:

$$F_{\text{sa}} = \sigma(f_{\text{conv}}([f_{\text{max}}(F_p), f_{\text{pool}}(F_p)])) \cdot F_p, \quad (6)$$

式中: σ 代表 sigmoid 函数, f_{conv} 代表核大小为 7×7 的卷积层函数, f_{max} 和 f_{pool} 分别代表最大池化层函数和平均池化层函数, F_{sa} 代表空间注意力机制模块的输出特征。

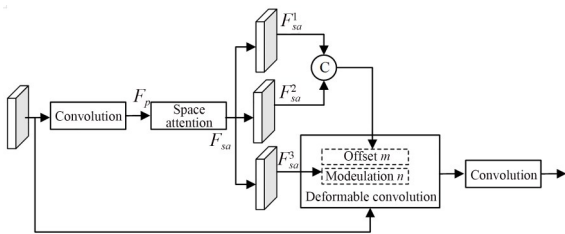


图 6 空间感知的可变形卷积的结构

Fig. 6 Architecture of space aware deformable convolution

在得到 F_{sa} 后,将它进一步划分为三个维度相同的特征 $F_{\text{sa}}^1, F_{\text{sa}}^2$ 和 F_{sa}^3 , 其中 F_{sa}^1 和 F_{sa}^2 串联作为 SADC 的偏移量 m , 而 F_{sa}^3 则作为 SADC 的调制值

n 。将 m 和 n 共同注入到 DC 的对应参数设置中, 获得 SADC, 即式(4)中 f_{sa} 。

2.1.2 纹理感知网络构建

本文使用 DTB 构建结构纹理提取网络, 其结构如图 7 所示。它共包含两个映射模块, 三对基于 DTB 的编码模块及下采样层, 一个基于 DTB 的特征细化层, 三对基于 DTB 的解码模块及上采样层。

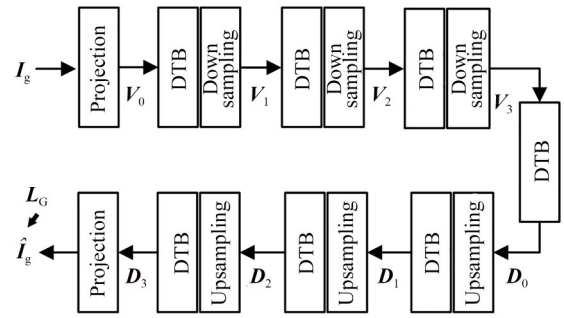


图 7 纹理感知网络结构

Fig. 7 Architecture of texture-aware network

对于给定初始水下图像的灰度图 $I_g \in \mathbb{R}^{l \times H \times W}$ (l 为通道数, H, W 为特征图的高度和宽度), 首先利用映射模块将其通道数扩充至 32, 得到维度为 $32 \times H \times W$ 的特征, 随后将该特征拉伸为 $HW \times 32$ 的向量 V_0 。再将向量 V_0 输入后续三个 DTB 的编码模块及上采样层中, 分别得到编码特征 V_1, V_2, V_3 , 其计算公式为:

$$V_0 = f_{\text{proj}}(I_g), \quad (7)$$

$$V_{i+1} = f_{\text{down}}(f_{\text{dtb}}^{\text{en}}(V_i)), \quad i = 0, 1, 2, \quad (8)$$

式中: f_{proj} 为映射层函数, 它由卷积层和 Reshape 层来实现维度为 $1 \times H \times W$ 的图像 I_g 到维度为 $32 \times H \times W$ 的特征 V_0 的映射; $f_{\text{dtb}}^{\text{en}}$ 为编码模块函数, 它由两个 DTB 组成; f_{down} 为下采样操作, 该操作将输入向量的通道数降为原先的四分之一, 向量长度上升为原先的 2 倍。因此, 当 i 分别取 0, 1, 2 时, 式(8)得到向量 V_1 (维度 $(HW/4) \times 64$), V_2 (维度 $(HW/16) \times 128$), V_3 (维度 $(HW/64) \times 256$)。将 V_3 进一步输入到基于 DTB 的特征细化层, 得到 $D_0 \in \mathbb{R}^{(HW/64) \times 256}$:

$$D_0 = f_{\text{dtb}}(V_3). \quad (9)$$

随后, 开始实施解码操作。与编码操作对应, 利用三个基于 DTB 的解码模块及上采样层

实施解码,其计算公式如下:

$$D_{j+1} = f_{\text{dbt}}^{\text{de}}(f_{\text{up}}(D_j)), j = 0, 1, 2, \quad (10)$$

式中: $f_{\text{dbt}}^{\text{de}}$ 为编码模块函数,它由2个DTB组成; f_{up} 为上采样操作,该操作与编码器中的下采样操作呼应,将输入向量的通道数扩充4倍,向量长度下降50%,最终获得维度为 $HW \times 32$ 的向量 D_3 。

将 D_3 输入映射层获得 $I \times H \times W$ 的输出图像 $\hat{I}_g$:

$$\hat{I}_g = f_{\text{proj}}(D_3), \quad (11)$$

式中: f_{proj} 的定义与式(7)中的定义保持一致。

为了使该网络能更有效地提取纹理结构特征,训练时使用真实图像的灰度图来监督 $\hat{I}_g$ 的更新:

$$L_G = \sum_{n=1}^N \|\hat{I}_g^n - G_g^n\|_2, n = 1, 2, \dots, N, \quad (12)$$

式中: L_G 为纹理结构提取网络的损失函数; $\hat{I}_g^n$, G_g^n , $n = 1, 2, \dots, N$ 为纹理结构提取网络的一系列输出结果及相应真实图像的灰度图像; $\|\cdot\|_2$ 为 L_2 损失函数。

2.2 颜色直方图特征提取网络

颜色直方图提取网络旨在从图像中提取不包含纹理结构的颜色特征。因此,利用高斯模糊滤波对初始水下图像 I 实施模糊操作,并将模糊后的结果 I_c 作为颜色信息提取网络的输入。与纹理结构提取网络使用的编码解码器作为网络框架不同,颜色直方图提取网络仅利用编码器作为网络框架,并用真实图像的颜色直方图监督编码器输出特征的更新。具体来说,颜色直方图提取

网络包含一个卷积层,三个残差模块(Residual Block, RB)以及三个下采样操作,其网络结构如图8所示。

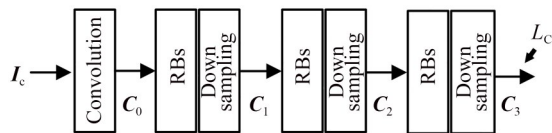


图8 颜色直方图信息提取网络

Fig. 8 Architecture of histogram color feature extraction network

对于给定的 I_c ,首先利用卷积层来提取通道数为32的浅层特征 C_0 :

$$C_0 = f_{\text{conv}}(I_c), \quad (13)$$

式中: f_{conv} 为 3×3 的卷积,它从 $I_c \in \mathbb{R}^{3 \times H \times W}$ 中提取维度为 $32 \times H \times W$ 的浅层特征 C_0 。

随后, C_0 进一步进入三个残差模块及下采样操作,对应函数为:

$$C_{i+1} = f_{\text{down}}(f_{\text{rb}}(C_i)), i = 0, 1, 2, \quad (14)$$

式中: f_{rb} 为残差模块函数,每个残差模块包含两个残差层,且每个残差层包含两个卷积和一个残差连接; f_{down} 为下采样操作函数,它将输入特征的通道数上升一倍,同时将特征尺度缩小一倍。由式(14)可见,当 $i = 0, 1, 2$ 时,获得维度大小分别为 $H/2 \times W/2 \times 64$, $H/4 \times W/4 \times 128$, $H/8 \times W/8 \times 256$ 的特征 C_1, C_2, C_3 。图9为颜色直方图提取网络在输入图2(b)后, C_0, C_1, C_2, C_3 的CAM可视化结果。

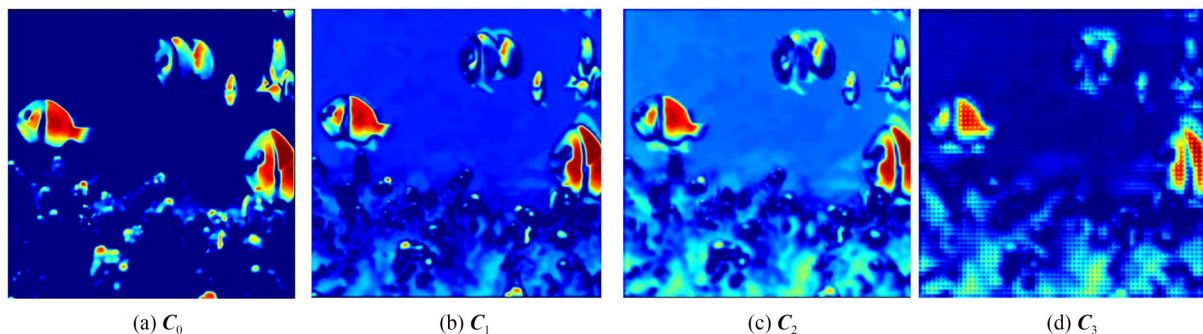


图9 不同维度的颜色特征图可视化结果

Fig. 9 Visualized results of color features with different dimensions

考虑到颜色直方图也是一种颜色特征,且此颜色特征中不包含图像的空间信息。将用真实

水下图像的三通道直方图向量 H (尺度为 $3 \times 256, 3$ 为颜色通道)来监督本网络颜色特征的提

取。将 C_3 输入到映射模块,得到维度为 3×256 向量:

$$A = f_{\text{prob}}(C_3), \quad (15)$$

式中: f_{prob} 为映射模块函数,该模块包含一个 3×3 的卷积层,一个池化层,一个 Reshape 层,一个线性层和一个 softmax 函数。 C_3 通过卷积层后,得到维度为 $H/32 \times W/32 \times (256 \times 3)$ 的特征。该特征进一步进入池化层,得到维度为 $1 \times 1 \times (256 \times 3)$ 的特征。经 reshape 层拉伸后,得到维度为 3×256 的向量。该向量顺序经过线性函数层和 softmax 函数层后,得到维度为 3×256 的向量 A 。图 10 和图 11 为本文模型在输入图 2(a) 后由向量 A 预测的直方图结果和真实图像直方图结果。利用真实图像的颜色直方图向量 H 监督 A 的更新,可以提取准确的颜色特征。

$$L_C = \sum_{n=1}^N \|A^n - H^n\|_1 / 3 \times 256, \quad n = 1, 2, \dots, N, \quad (16)$$

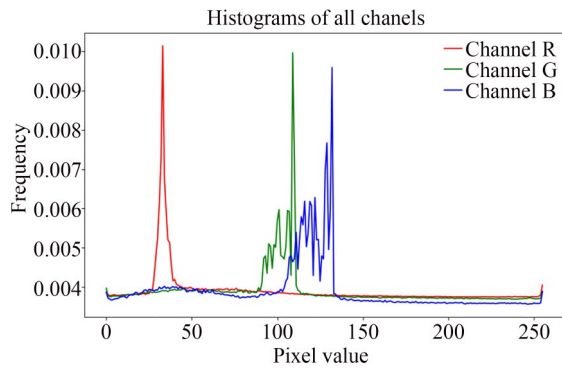


图 10 预测直方图结果

Fig. 10 Predicted histogram results

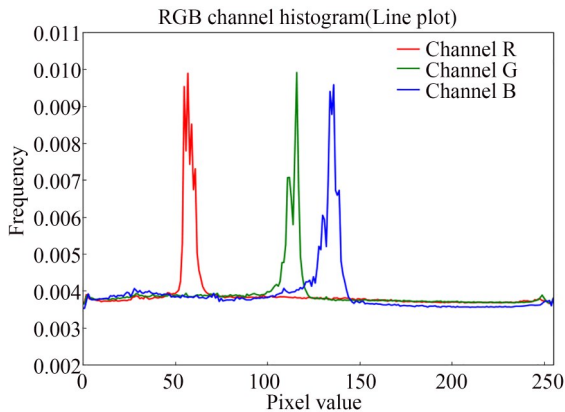


图 11 真实直方图结果

Fig. 11 Real histogram results

式中: L_C 为颜色信息提取网络的损失函数; A^n , H^n , $n = 1, 2, \dots, N$ 为颜色直方图特征提取网络的一系列输出向量及相应真实图像的直方图; $\|\cdot\|_1$ 为 L_1 损失函数。

2.3 颜色纹理融合的水下图像增强网络

颜色纹理融合的水下图像增强网络进一步利用所提的颜色纹理融合模块(见图 1 中 CTF_i 模块),将前两个网络提取的纹理向量 V_1, V_2, V_3 与颜色特征 C_1, C_2, C_3 有效融合,用于水下图像增强,该操作不但能使增强后的图像具有清晰的纹理细节和真实的颜色,还能确保颜色和纹理结构的一致性。

2.3.1 颜色纹理融合模块

由图 1 可见,颜色纹理融合模块(CTF_i 模块)用于将纹理向量 V_i 与颜色特征 C_i 融合。本网络共有 3 个颜色纹理融合模块 $CTF_1, CT F_2$ 和 $CT F_3$,分别融合 V_1 (维度 $(HW/4) \times 64$)与 C_1 (维度 $H/2 \times W/2 \times 64$), V_2 (维度 $(HW/16) \times 128$)与 C_2 (维度 $H/4 \times W/4 \times 128$), V_3 (维度 $(HW/64) \times 256$)与 C_3 (维度 $H/8 \times W/8 \times 256$)。考虑到纹理向量 V_i 和颜色特征 C_i 中信息的差异性,采用串联操作融合会导致增强结果失真。因此,颜色纹理融合模块利用颜色纹理关系矩阵 R_i 为权重,实施融合。

如图 12 所示,颜色纹理融合模块首先将向量 V_i 拉伸为与 C_i 维度一致的特征 V'_i ,随后采用曼哈顿距离和元素内积的方式,获得颜色和纹理的

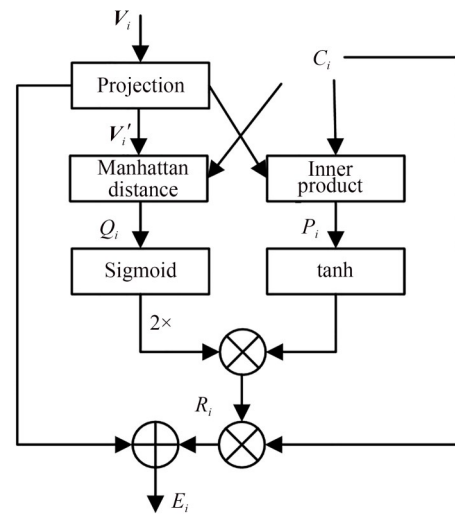


图 12 颜色纹理融合模块结构

Fig. 12 Architecture of color texture fusion module

关系矩阵 R_i , 计算过程为:

$$V'_i = f_{\text{proj}}(V_i), i = 1, 2, 3, \quad (17)$$

$$Q_i(x, y) = -\|V'_i(x, y) - C_i(x, y)\|_1, i = 1, 2, 3, \quad (18)$$

$$P_i(x, y) = V'_i(x, y) \times C_i(x, y), i = 1, 2, 3, \quad (19)$$

$$R_i = 2 \times \lambda(Q_i) \odot \tau(P_i), i = 1, 2, 3, \quad (20)$$

式中: f_{proj} 的定义与式(5)中保持一致; V'_i 为映射后的特征, 映射后 V'_i 的维度与 C_i 维度一致; (x, y) 为特征的元素位置索引; $\|\cdot\|_1$ 为 L_1 范数; λ 为 sigmoid 函数, τ 为 tanh 函数; $\odot$ 代表点乘。

式(20)得到的关系矩阵 R_i 准确计算了特征 V'_i 与 C_i 的关系, 再实施以下融合, 得到融合特征 E_i :

$$E_i = R_i \odot C_i + V'_i, i = 1, 2, 3, \quad (21)$$

式中: E_1 维度为 $H/2 \times W/2 \times 64$, E_2 的维度为 $H/4 \times W/4 \times 128$, E_3 的维度为 $H/8 \times W/8 \times 256$ 。由图 1 可见, 融合后的特征 E_i 会进一步注入到后续的颜色纹理融合水下图像增强网络。

2.3.2 颜色纹理融合的水下图像增强网络构建

进一步将 E_i 其嵌入基于编码解码框架的网络中, 以构建基于颜色纹理融合的水下图像增强网络。这里, 编码解码框架包含一个卷积层, 三个残差模块及相应的下采样操作, 三个残差模块及相应的上采样操作, 一个重构层。网络结构如图 13 所示。

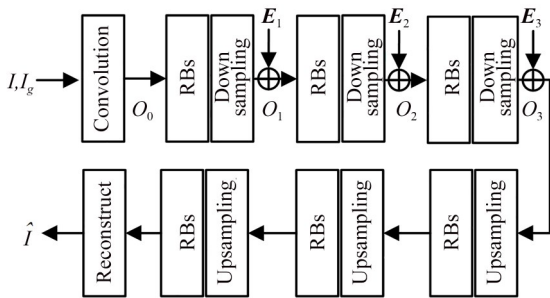


图 13 颜色纹理融合的水下图像增强网络结构

Fig. 13 Architecture of color-texture fusion-based underwater image enhancement network

首先, 串联初始水下图像 I 及纹理感知网络输出的图像 $\hat{I}_g$, 并输入到一个卷积层中, 即:

$$O_0 = f_{\text{conv}}([I, \hat{I}_g]), \quad (22)$$

式中: f_{conv} 的定义与式(13)一致; $[\]$ 为串联操作; O_0 为卷积层的输出。

若将 O_0 直接输入到后续的残差模块及下采样操作中, 则对应的函数为:

$$O_{t+1} = f_{\text{down}}(f_{\text{rb}}(O_t)), t = 0, 1, 2, \quad (23)$$

式中: f_{rb} 和 f_{down} 的定义和设计式(14)中保持一致; f_{down} 为下采样操作函数, 当 $t = 0, 1, 2$ 时, 获得维度分别为 $H/2 \times W/2 \times 64$, $H/4 \times W/4 \times 128$, $H/8 \times W/8 \times 256$ 特征 O_1, O_2, O_3 。

然而, 在实际操作中, 为了将颜色纹理融合特征 E_1, E_2, E_3 有效地嵌入到水下图像增强网络, 将维度相同的 O_1, O_2, O_3 与 E_1, E_2, E_3 相加后输入到后续的残差模块及下采样操作。因此, $t = 1, 2$ 时, 执行以下计算:

$$O_{t+1} = f_{\text{down}}(f_{\text{rb}}(O_t + E_t)), t = 1, 2. \quad (24)$$

在得到 O_3 后, 将 $O_3 + E_3$ 的结果进一步输入到后续的解码器中:

$$\hat{I} = f_{\text{re}}(f_{\text{rb}}(f_{\text{up}}(f_{\text{rb}}(f_{\text{up}}(f_{\text{rb}}(f_{\text{up}}(O_3 + E_3))))))), \quad (25)$$

式中: f_{up} 为上采样操作函数, 由反卷积实现, 它将输入特征的通道数下降 50%, 尺度上升一倍; f_{rb} 的定义与式(14)中保持一致; f_{re} 为重构层函数, 由卷积层来实现, 它将特征重构为增强后的水下图像 $\hat{I}$ 。

2.4 损失函数

选用 L_2 作为损失函数来最小化颜色纹理融合的水下图像增强网络输出结果 $\hat{I}$ 与真实结果 G 的差异。假设 $\hat{I}_n$ 和 $G_n, n = 1, 2, \dots, N$ 分别为一系列模型输出的增强结果及对应的真实结果, 则损失函数 L_2 计算如下:

$$L_2 = \frac{\sum_{n=0}^N \|\hat{I}_n - G_n\|_2}{N}. \quad (26)$$

这里利用 L_G 监督纹理感知网络, 提取纹理特征, 利用 L_C 监督颜色直方图特征提取网络提取颜色特征。纹理感知网络和颜色直方图特征提取网络, 又通过颜色纹理融合模块融合将其与后续的水下图像增强网络关联。因此, 在模型训练过程中, 将联合 L_G, L_C 和 L_2 作为整体模型的损失函数 L :

$$L = L_2 + L_G + L_C. \quad (27)$$

3 实验结果与分析

3.1 数据集及实施细节

为了确保模型的泛化能力,选用的数据集^[7]包含 10 种不同水类型(I, IA, IB, II, III, 1, 3, 5, 7, 9 水类型)的合成图像,其中 I, IA, IB, II, III 为公海水域类型,1, 3, 5, 7, 9 为沿海水域类型。具体来说,文献[7]使物理水下成像模型(式(28))对 1 449 张 NYU-V2 室内图像实施合成。

$I_\lambda(x) = J_\lambda(x)t_\lambda(x) + A_\lambda(x)(1 - t_\lambda(x))$, (28)
式中: λ 为红绿蓝光的波长, I_λ 是观测到水下图像, J_λ 是清晰图像(合成是为室内图像), A_λ 为全局背景光, t_λ 为传输系数,描述了光从水中到达相机时,未被水吸收或散射的光比例。在合成过程中,全局背景光 A_λ 从 $[0.8, 1]$ 中随机选取, t_λ 的选取依据如下:

$$t_\lambda(x) = e^{-\beta^\lambda d(x)} = \frac{E_\lambda(x, d(x))}{E_\lambda(x, 0)} = N_\lambda(d(x)), \quad (29)$$

式中: $E_\lambda(x, 0)$ 代表光在点 x 处的初始能量, $E_\lambda(x, d(x))$ 代表光传播 $d(x)$ 距离后的剩余能量, N_λ 是剩余能量与初始能量的比率,并随着 $d(x)$ 的变化而变化。在 0.5~15 m 随机选择 $d(x)$ 值, N_λ 对不同水类型不同红绿蓝光波长的取值如表 1 所示。合成过程中,对每一张图像,使用不同的 A_λ 和 t_λ 可合成 10 张不同水类型图像,因此,1 449 张 NYU-V2 室内图像共合成 14 490 张训练图像,而初始的 NYU-V2 室内图像则为真实结果。从中选取 13 041 张图像为训练集,1 449 张为测试集。此外,还采用两个真实图像数据集用于性能测试:(1)真实水下图像增强数据集(Real Underwater Image Enhancement Benchmark, UIEB)^[7],该数据集包含 890 张多样化的真实水下图像,且每张图像均有增强结果参考图,故可用于算法性能的量化评估;(2)挑战数据集^[7],包含了 60 张降质严重,低光照的挑战性水下图像。

表 1 10 种水类型下不同光波长的 N_λ 值

Tab. 1 Values of N_λ under different water types and light wavelengths

Type	R	G	B	Type	R	G	B
I	0.982	0.961	0.805	1	0.875	0.885	0.750
IA	0.975	0.955	0.804	3	0.800	0.820	0.710
IB	0.968	0.950	0.830	5	0.670	0.730	0.670
II	0.940	0.925	0.800	7	0.500	0.610	0.620
III	0.890	0.885	0.750	8	0.290	0.460	0.550

本文所提的网络在 NVIDIA RTX 2080 Ti GPU 的 PC 机上,通过搭建 Pytorch 深度学习框架实现。优化算法 Adam 的学习率为 0.000 1, β_1, β_2 分别取默认值 0.9, 0.999。训练过程中,所有网络的输入图像尺寸 $H \times W$ 均为 512×512 ,其中,纹理感知网络的输入为水下图像的灰度图 I_g ,颜色直方图特征提取网络的输入为水下图像的高斯模糊图 I_c ,颜色纹理融合的水下图像增强网络输入为水下图像 I 及纹理感知网络输出的图像 $\hat{I}_g$ 。使用式(27)作为损失函数。训练次数为 400 epochs。

3.2 量化指标

本文选取 6 个常用的量化指标,分别为 4 个全参考的图像质量评价指标:信噪比(Peak Sig-

nal to Noise Ratio, PSNR), 结构相似性(Structural Similarity Index Measure, SSIM), 局部批次对比度(Patch-based Contrast Quality Index, PC-QI), 以及 2 个无参考图像质量评价指标:水下图像质量测量指标(Underwater Image Quality Metric, UIQM)和信息熵(Entropy)。其中,PSNR 用来衡量两个图像之间差异的指标,其计算公式如下:

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right), \quad (30)$$

式中:MAX 为图像中像素的最大可能值,在 8 位的位图中通常为 255, MSE 为增强结果与真实结果的均方误差。

SSIM 从亮度、对比度和结构三个方面测量

两个图像之间的相似度,其计算公式如下:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + \alpha_1)(2\sigma_{xy} + \alpha_2)}{(\mu_x^2 + \mu_y^2 + \alpha_1)(\sigma_x^2 + \sigma_y^2 + \alpha_2)}, \quad (31)$$

式中: x, y 分别为增强结果与真实结果, μ_x, μ_y 为图像 x, y 的平均值, σ_x^2, σ_y^2 为图像 x, y 的方差, σ_{xy} 是图像 x, y 的协方差, α_1, α_2 为两个常数,用于避免分母中的零。

PCQI用于计算局部区域的对比度,其计算公式如下:

$$PCQI(x, y) = \frac{1}{M} \sum_{i=1}^M \omega_i PCQI_{patch}(x_i, y_i), \quad (32)$$

式中: x, y 分别为增强结果与真实结果, M 为图像被划分的局部块数, x_i, y_i 为第 i 个局部块的对应区域, ω_i 为对应区域的权重。 $PCQI_{patch}(x_i, y_i)$ 是基于局部块的对比度,亮度和结构等信息设置的局部块对比度质量指标。

UIQM通过统计图像的色彩,对比度和边缘锐度的线性组合来衡量增强图像的质量,计算公式为:

$$UIQM(x, y) = \frac{1}{L} \sum_{l=1}^L \omega_l \frac{4\sigma_{xy}^2}{(\sigma_x^2 + \sigma_y^2)(\mu_x^2 + \mu_y^2)}, \quad (33)$$

式中: x, y 分别为增强结果与真实结果; L 为图像被划分的局部块数; σ_{xy}^2 是共变方差; σ_x^2 和 σ_y^2 分别是 x, y 的方差; μ_x^2, μ_y^2 分别为 x, y 的平均值平方, ω_l 为每个局部块的权重。

Entropy量化图像包含信息量,计算公式如下:

$$Entropy = - \sum_{i=1}^N P(q_i) \log_2(P(q_i)), \quad (34)$$

式中: N 为图像中所有灰度级别数量; $P(q_i)$ 为灰度级别 q_i 出现的概率。PSNR, SSIM, PCQI, UIQM和Entropy的值越大,增强图像的视觉效果越好。

3.3 消融实验

3.3.1 网络框架的有效性验证

本模型的网络框架主要由纹理感知网络,颜色直方图特征提取网络和水下图像增强网络组成。其中,纹理感知网络和颜色直方图特征提取网络为后续水下图像增强网络提供纹理和颜色特征。为了验证纹理和颜色特征对水下图像增

强的有效性,设计了两个变异模型:w/o G和w/o C。其中,w/o G从所提网络模型中移除纹理感知网络以及相应的纹理颜色融合模块,将颜色直方图特征网络提取的信息直接加到水下图像增强网络的编码特征中;w/o C代表从所提网络模型中移除颜色直方图特征提取网络和纹理颜色融合模块,将纹理感知网络提取的纹理特征直接加到水下图像增强网络的编码特征中。表2呈现的定量结果表明,当纹理感知网络和颜色直方图特征提取网络被移除后,所提出的变异模型性能显著下降。上述实验结果进一步证实了将颜色特征与纹理信息相结合,能共同实现水下图像增强的合理性和有效性。

表2 合成图像集上的消融实验结果

Tab. 2 Results of ablation studies on synthetic images

Model	PSNR	SSIM
w/o G	22.15	0.853 5
w/o C	22.23	0.862 2
Depth-MSA	24.11	0.899 1
Texture-RB	23.74	0.882 9
w/o F	23.65	0.880 7
Ours	25.35	0.903 4

3.3.2 网络模块的有效性验证

所提模型采用DTB构建纹理感知网络,提取纹理特征。这里,多头注意力机制是DTB的关键,它利用空间感知的变形卷积层SADC来计算多头注意力机制的 Q, K, V 值,使得DTB具有几何感知能力与全局语义捕获能力。为验证DTB的有效性,设计了两个变异模型:Depth-MSA和Texture-RB。Depth-MSA利用深度卷积代替SADC计算 Q, K, V 值。而Texture-RB则将图7中的DTB模块换成了2.2节的残差模块(Residual Block, RB)。在合成图像上的测试结果如表2所示,本文算法性能明显优于Depth-MSA和Texture-RB。为了进一步证明DTB具有较好的纹理感知和全局语义结合能力,图14分别列出了本文模型Depth-MSA和Texture-RB输入图2(a)后提取的 V_3 特征可视化结果。从中可见,本文算法因将利用变形卷积来计算多头注意力机制,不但能感受全局语义即图像中包含的珊瑚和鱼,还能关注鱼和珊瑚的细小纹理,见图14(a)

中箭头所示。这使得增强结果能锐化相关区域,获得较好的增强结果。Depth-MSA 模型利用深度卷积来计算多头注意力机制,虽能仅能感知全局语义,如鱼及珊瑚,但对于细小的珊瑚细节和鱼的纹理则无法有效感知,如 14(b)所示。相比之下,Texture-RB 仅用 RB 来构建网络,既不能感知全局语义也不能感知纹理,仅能从输入的图片中捕获显著性最大的纹理细节,如图 14(c)所示,颜色较亮的关注区域仅为珊瑚顶上的边缘。由此表明,基于 SADC 的多头注意力机制具有更好

的纹理感知能力,能帮助网络更好地增强水下图像。

另一方面,颜色纹理融合模块利用颜色纹理关系矩阵将提取的纹理特征与颜色特征有效融合,能确保颜色和纹理一致性。为了验证该模块的有效性,设计了 w/o F 变异模型,该模型在所提模型的基础上,删除颜色纹理融合模块,并将颜色和纹理特征直接加在编码特征上,对应的结果如表 2 所示。从中可见, w/o F 的性能明显下降,由此证明颜色纹理融合模块的贡献。

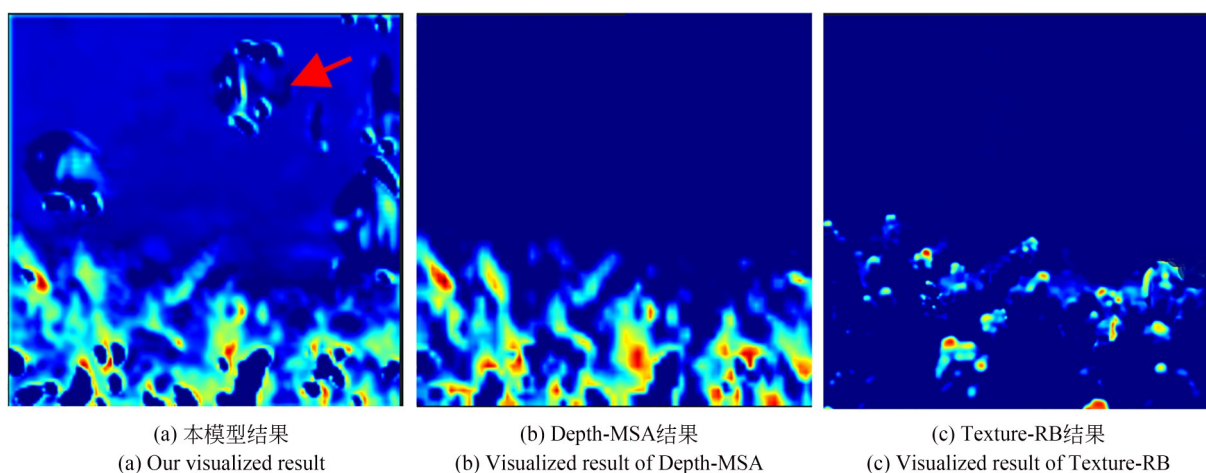


图 14 不同变异模型的特征可视化结果

Fig. 14 Visualized results of different variant models

3.4 合成水下图像增强结果对比

从合成水下图像集挑选了 3 张极具挑战性的合成图像进行不同算法的视觉效果对比,结果如图 15 所示。对比算法选用传统方法 HUIE^[4],以及基于深度学习算法 WDN^[9],ALAN^[10],Swin-conv^[11]和 U-T^[12]。由图 15(f)所示,U-T 的颜色校正结果不理想,增强结果中引入了明显的黄绿色。HUIE 能有效去除原图像的蓝绿色偏,但过度增强使图像颜色失真,图像过亮,如图 15(b)所示。ALAN 和 Swin-conv 算法的色彩恢复性较差,增强的结果出现了亮度偏低及雾化的现象,如图 15(d)和 15(e)所示。WDN 的增强结果颜色偏暗,如图 15(c)所示。本文算法的增强结果色彩真实,细节丰富,与真实参考图像最接近。每幅图像下方标注的 PSNR/SSIM 值也表明,本文算法具有最高的 PSNR 及 SSIM 值。为了进一步验证本文算法的有效性,采用 PSNR,SSIM,

EI,PCQI,UIQM 和 Entropy 在合成图像测试集上进行了量化评估,各指标平均量化结果如表 2 所示。由表可见,本文算法仍具有最高的量化指标值。

3.5 真实水下图增强结果对比

在真实图像数据集 UIEB 上测试各算法的水下图像增强效果。从 UIEB 中挑选了 3 幅挑战性的水下图像(见图 16(a))。各算法对 3 幅图像的增强结果如图 16(b)~16(g)所示。从中可见, HUIE 与 Swin-conv 算法不但没有消除色偏,还出现了颜色失真,其增强结果呈红色调及紫色调,如图 16(b)和 16(e)的第三行所示。WDN 算法在蓝绿水下图像上过度增强,视觉结果出现亮度不均匀,局部过暗的现象,如图 16(c)第三行的背景。U-T 算法的增强结果出现雾化现象,图像细节模糊,如图 16(f)所示。ALAN 算法的增强结果对比度过大,色调失真,如图 16(d)第二行所

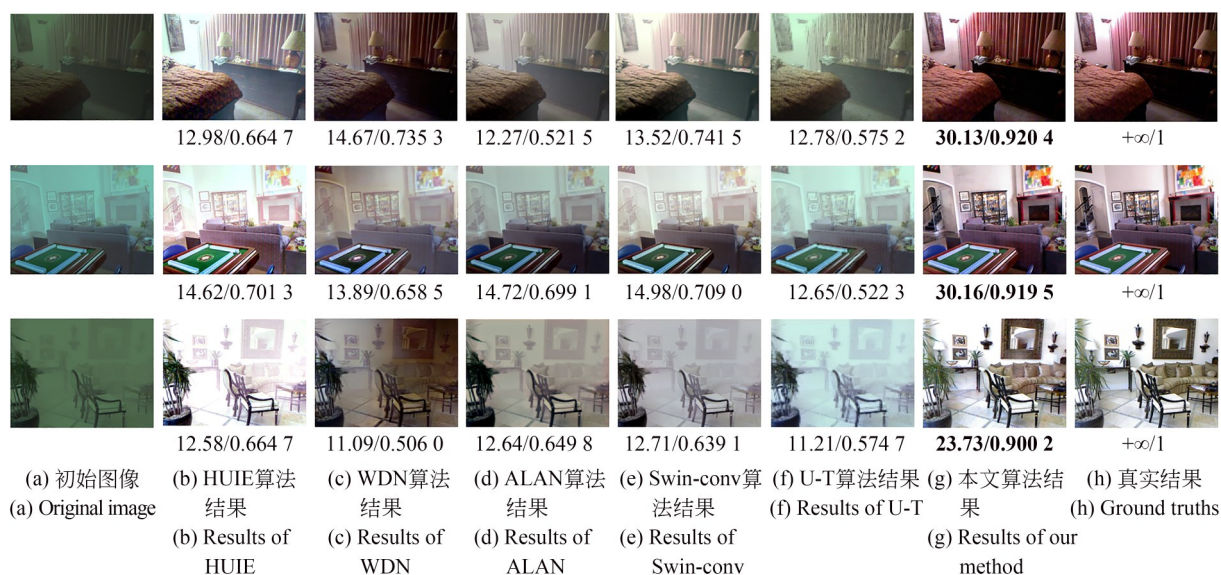


图 15 合成水下图像增强结果对比

Fig. 15 Comparison of enhanced results on synthetic underwater images

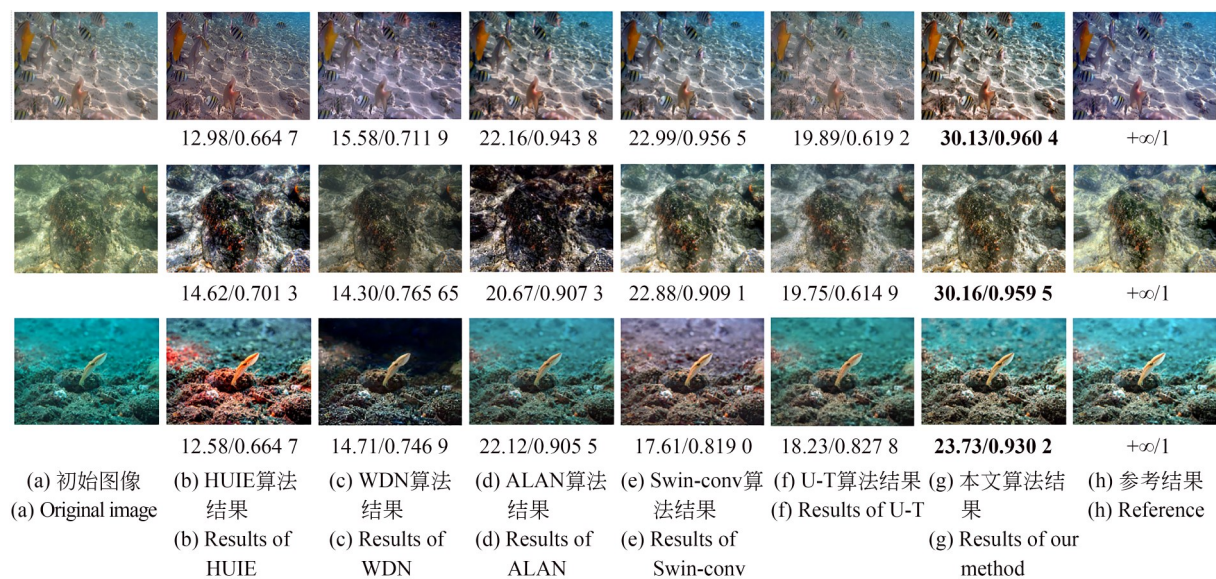


图 16 真实水下图像增强结果对比

Fig. 16 Comparison of enhanced results on real underwater images

示。本文算法不但能有效校正色偏,保留纹理细节,还在一定程度上提高了图像的对比度和亮度(见图 16(g)),增强结果甚至比参考结果(见图 16(h))更真实、自然。由于 UIEB 测试集提供了参考图像,除选用无参考的量化指标 UIQM, Entropy 来量化视觉结果外,还选用全参考的量化指标 PSNR, SSIM, EI, PCQI 实施视觉效果评价。各算法在 UIEB 上的平均量化结果如表 3 所示。

由表 3 可见,本文算法在 PSNR, SSIM, EI,

PCQI, UIQM 和 Entropy 上均获得最高值,说明本文算法在纹理,颜色,结构信息方面有更好的视觉增强效果。

3.6 挑战性真实水下图像增强结果对比

从 60 张挑战性图像上选取 2 张弱光照的水下图像(见图 17(a)),各算法的增强结果如图 17(b)~17(g)所示。由图可见,ALAN 无法有效地校正色偏,其增强结果引入黄色的色偏,如图 17(d)第一行的背景呈黄色调。WDN, Swin-

表 3 不同算法在合成和真实图像测试集上的结果对比

Tab. 3 Comparison of enhanced results by different algorithms on synthetic and real image datasets

Model	Synthetic images						Real images					
	PSNR	SSIM	EI	PCQI	UIQM	Entropy	PSNR	SSIM	EI	PCQI	UIQM	Entropy
HUIE ^[4]	11.48	0.635	66.52	0.968	2.395	7.029	18.51	0.803 0	74.89	0.986	2.827	7.338
WDN ^[9]	15.88	0.725	71.97	0.971	2.556	7.123	15.23	0.676 3	68.45	0.963	2.454	7.022
ALAN ^[10]	15.52	0.731 1	72.19	0.970	2.591	7.194	19.25	0.871 8	73.89	0.977	2.840	7.417
Swin-conv ^[11]	15.48	0.737 7	74.55	0.973	2.751	7.241	23.37	0.890 7	82.58	0.994	2.992	7.553
U-T ^[12]	14.78	0.575 3	81.52	1.054	3.111	7.824	20.67	0.880 6	81.45	1.037	2.986	7.527
Ours	25.28	0.903 4	83.67	1.105	3.276	7.952	24.63	0.900 6	83.55	1.096	3.105	7.658

conv,U-T 无法对低照明区域实施视觉增强,增强后图像细节模糊,如图 17(c),17(e)和 17(f)第二行的砂石和鱼丢失细节。HUIE 的增强结果清晰度较高,但整体颜色偏暗,如图 17(b)所示。相比而言,本文算法不但能有效去除色偏,提升视

觉对比度,还能保留细节,其增强结果(见图 17(g))具有最佳的视觉效果。表 4 进一步给出各算法在 60 张挑战性水下图像集上的平均 UIQM 和 Entropy 值,从中可见,本文算法的量化指标最高,增强效果最优。

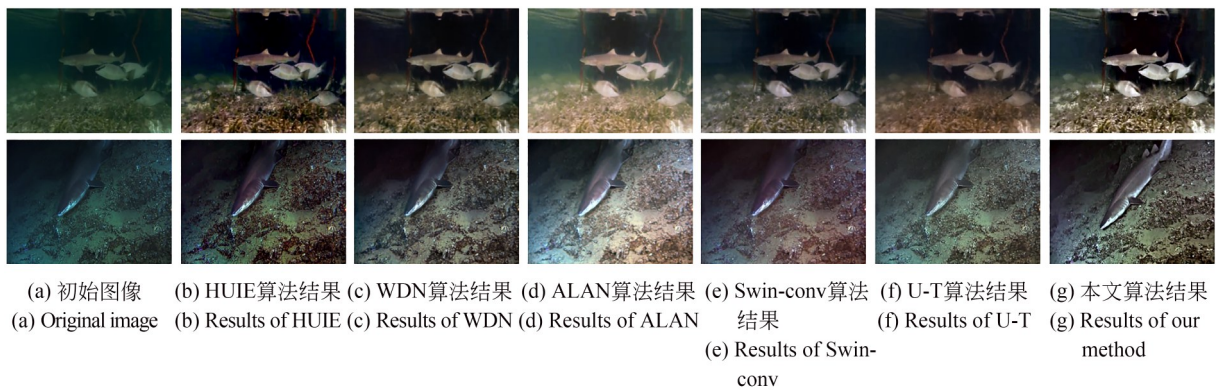


图 17 挑战性水下图像增强结果对比

Fig. 17 Comparison of enhanced results on challenging underwater images dataset

表 4 不同算法在挑战性图像集的结果对比

Tab. 4 Comparison of different methods on challenge and UCCS datasets

Model	Challenge	
	UIQM	Entropy
HUIE	2.341	5.229
WDN	2.454	6.110
ALAN	2.598	6.654
Swin-conv	2.698	7.012
U-T	2.577	6.741
Ours	2.897	7.212

3.7 运行时间对比

本网络框架包含 3 个子网络。为测试运行时间,在包含 1 449 张,尺寸为 512×512 的合成图像测试集上进行本算法与其他水下增强算法的平均运行时间对比。所有算法均在相同的计算机上运行,包括:在 Matlab 平台上运行的传统算法 HUIE^[4],在 Pycharm 平台上运行的深度学习算法 WDN,ALAN,Swin-conv,U-T,以及本文算法。表 5 为各算法的平均运行时间对比。从表中可见,本文算法位于第二,单张水下图像的处理时间明显优于传统算法及基于深度学习的水下图像增强算法 ALAN,Swin-conv 和 U-T。WDN 算法因由两个简单的编码解码网络组成,

网络更轻,运行时间最短,仅需要 0.032 s。不过,本文算法虽具有次优的运行时间 0.051 s,较其他水下图像增强算法的运行时间至少缩短了 9%,且视觉增强效果最好(见表 3),可用于水下视觉增强任务中。

表 5 不同算法的平均运行时间对比

Tab. 5 Running time comparison of different algorithms (s)

Model	Time
HUIE	2.351
WDN	0.032
ALAN	0.056
Swin-conv	0.173
U-T	0.183
Ours	0.051

4 结 论

本文以水下图像增强为研究对象,提出了纹理感知联合颜色直方图特征的水下图像增强网络,它包含纹理感知网络、颜色直方图特征提取网络以及颜色纹理融合的水下图像增强网络。首先,本文采用纹理感知网络来感知纹理,它从

灰度图中提取纹理特征。这里采用所提的变形 Transformer 模块来构建纹理感知网络,该模块利用空间感知的变形卷积来计算 Transformer 模块中多头注意力机制中 K, Q, V 值,从而有效地将变形卷积的几何感知能力与 Transformer 模块的全局语义捕获能力相结合,提升模块的纹理结构感知力。随后,利用编码器来构建颜色直方图特征提取网络。该网络为提取不包含纹理的颜色特征,将高斯模糊的水下图像作为此网络的输入,并利用真实水下图像的直方图来监督该网络颜色特征的提取。最后,利用所提的颜色纹理融合模块计算前两个网络提取的颜色和纹理特征的关系矩阵,以有效融合颜色和纹理特征,用于颜色纹理融合的水下图像增强网络实施增强,该操作不但有效地保留了纹理结构,校正了失真的颜色,还使得纹理和颜色信息保存一致。实验结果表明,本文算法较存在算法具有更好的视觉增强效果,量化指标 UIQM 至少提高了 10%,运行时间缩短了 9%,仅 0.051 s,能满足水下机器视觉的要求。在后续工作中,计划开展面向真实场景的无监督水下图像增强网络,并引入场景识别子网络,以预先识别水下图像的环境类型,以提高算法的泛化能力。

参考文献:

- [1] 王昊天,刘庆省,陈亮,等.改进的 CycleGAN 网络用于水下显微图像颜色校正[J]. 光学精密工程, 2022, 30(12): 1499-1508.
WANG H T, LIU Q SH, CHEN L, *et al.* Improved CycleGAN network for underwater microscopic image color correction [J]. *Opt. Precision Eng.*, 2022, 30(12): 1499-1508. (in Chinese)
- [2] DREWS P L J Jr, 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.
- [3] SONG W, WANG Y, HUANG D M, *et al.* Enhancement of underwater images with statistical model of background light and optimization of transmission map[J]. *IEEE Transactions on Broadcasting*, 2020, 66(1): 153-169.
- [4] LI X J, HOU G J, TAN L, *et al.* A hybrid framework for underwater image enhancement[J]. *IEEE Access*, 2020, 8: 197448-197462.
- [5] 胡振宇,陈琦,朱大奇.基于颜色平衡和多尺度融合的水下图像增强[J]. 光学精密工程, 2022, 30(17): 2133-2146.
HU ZH Y, CHEN Q, ZHU D Q. Underwater image enhancement based on color balance and multi-scale fusion [J]. *Opt. Precision Eng.*, 2022, 30(17): 2133-2146. (in Chinese)
- [6] ANWAR S, LI C, PORKLI F. Deep underwater image enhancement [J]. *arXiv preprint arXiv: 1807.03528*, 2018.
- [7] 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. DOI: 10.1109/TIP.2019.2955241.
- [8] WANG N, ZHOU Y B, HAN F, *et al.* UW-GAN: underwater GAN for real-world underwater color restoration and dehazing [J]. *arXiv preprint*

- arXiv:1912.10269*, 2019
- [9] MA Z Y, OH C. A wavelet-based dual-stream network for underwater image enhancement [C]. *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. May 23-27, 2022. Singapore, Singapore. IEEE, 2022: 2769-2773.
- [10] 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.
- [11] REN T D, XU H Y, JIANG G Y, *et al.* Reinforced swin-convs transformer for simultaneous underwater sensing scene image enhancement and super-resolution [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 3205061.
- [12] PENG L T, ZHU C L, BIAN L H. U-shape transformer for underwater image enhancement [J]. *IEEE Transactions on Image Processing*, 2023, 32: 3066-3079.

作者简介:



袁国铭(1978—),男,四川达州人,博士,讲师,2012年于中国石油大学获得博士学位,主要从事数据图像处理,大数据分析,应急管理信息化的研究。
E-mail: scdyuan@126.com