

DOI: 10.13382/j.jemi.B2508202

基于关键特征传递的 KFT-GAN 水下图像增强模型*

麦仁贵¹ 王 骥^{2,3}

(1. 广东海洋大学数学与计算机学院 湛江 524088; 2. 广东海洋大学电子与信息工程学院 湛江 524088;
3. 广东省智慧海洋传感网及其装备工程技术研究中心 湛江 524088)

摘要:针对水体对光的吸收和散射等现象导致的水下图像出现色偏、模糊和细节丢失等问题,提出了基于关键特征传递的 KFT-GAN 水下图像增强模型。设计了 KFT 关键特征传递模块,解决了色彩、边缘和纹理等关键特征在网络中高效传递的问题,并结合深度可分离卷积构建了轻量化的网络模型,与使用普通卷积构建的模型相比,参数量降低了 60.6%,增强了模型的学习效率。同时该模块有利于生成器网络的编码阶段从输入图像中获取充足的关键特征,并通过下采样和跳跃连接将提取的关键特征传递至解码阶段进行图像重建,增强了重建图像的质量。另外,基于感知损失原理提出了混合损失函数,同时强调图像的多个关键属性,获得了视觉质量更好的图像效果。模型在 EUVP 和 UIEB 数据集上均取得了较好的性能,PSNR 值分别为 21.384 2 和 18.025 6,SSIM 值分别为 0.741 3 和 0.688 9,通过与传统算法和深度学习算法进行定性和定量的对比实验都证明了该模型的有效性和优越性。

关键词: KFT-GAN; 关键特征传递模块; 深度可分离卷积; 混合损失函数

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

Underwater image enhancement model based on generative adversarial network with key feature transfer

Mai Rengui¹ Wang Ji^{2,3}

(1. College of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang 524088, China;
2. College of Electronics and Information Engineering, Guangdong Ocean University, Zhanjiang 524088, China;
3. Research Center of Guangdong Smart Oceans Sensor Networks and Equipment Engineering, Zhanjiang 524088, China)

Abstract: To address the challenges of color cast, blurriness, and detail loss in underwater images caused by light absorption and scattering, this study proposes a KFT-GAN for underwater image enhancement. The paper introduces a KFT module that facilitates the efficient transmission of critical features such as color, edges, and texture within the network. By integrating depthwise separable convolution, a lightweight network model is constructed, reducing the parameter count by 60.6% compared to models utilizing conventional convolution, thereby enhancing the model's learning efficiency. This module enables the generator's encoding stage to extract essential key features from the input image and transmit these features to the decoding stage through downsampling and skip connections, improving the quality of the reconstructed image. Additionally, this study proposes a hybrid loss function based on perceptual loss principles, emphasizing multiple key attributes of the image to achieve superior visual quality. The proposed model demonstrates excellent performance on both the EUVP and UIEB datasets, achieving PSNR values of 21.384 2 and 18.025 6, and SSIM values of 0.741 3 and 0.688 9, respectively. Qualitative and quantitative comparisons with traditional algorithms and deep learning methods validate the effectiveness and superiority of the proposed model.

Keywords: KFT-GAN; key feature transfer module; depthwise separable convolutions; hybrid loss function

0 引言

海洋图像信息资源对研究海洋生物和勘探资源有重要意义。然而光在水下传播过程中受到水介质的吸收和散射的影响以及水下复杂的成像环境,会导致水下图像出现细节模糊、色度失真与对比度低等现象^[1]。对进一步图像分割、图像深度估计^[2]等高级视觉分析任务带来阻碍。因此,机器视觉研究领域学者对水下退化图像进行增强与恢复的研究非常关注。

Hu 等^[3]将水下图像增强方法分为传统方法和基于深度学习的方法;Rao 等^[4]从光谱分布的角度分析水下数据集,发现与非水下数据集相比,增强的参考图像的声音光谱分布更明显,所以提出了一种概率颜色补偿网络,通过纹理和颜色特征的多尺度体积融合估计颜色的概率分布。但是如果输入数据中存在噪声或伪影,可能会影响网络的效果,导致颜色补偿的准确性下降;Chang 等^[5]提出了一种概率网络来学习退化水下图像的增强分布,使用 Transformer 模型捕获远程像素相互作用,通过条件变分自编码器和自适应实例归一化构建增强分布,再根据分布中的一组样本预测确定性结果,但是 Transformer 模型计算复杂度较高,增加了训练过程的难度,依赖于大量高质量的训练数据来准确建模水下图像的退化和增强分布;Naik 等^[6]使用 3 个紧密连接的卷积块串联组成一个全连通的卷积网络,并使用跳过连接将输入图像连接到每个块的输出,达到图像增强的效果;Alyasser 等^[7]基于提出了基于冠状病毒群体免疫优化算法的水下图像增强方法 CHIO-UIE,改善了低对比度和光照下的水下图像;Yang 等^[8]使用卷积和反卷积结构构建了一个无参考图像的水下图像增强网络,通过跳跃连接将输入图像得到的无参考损失加入到网络输出的图像中,增强了水下退化图像的色彩和对比度;Chen 等^[9]提出了基于 CNN 网络的最优水下图像增强复合骨干网,满足一定的水下图像增强效果和实时性要求;Xu 等^[10]结合大核注意力模块和动态增强模块提出了基于交叉注意力的 U-Net 水下图像增强模型,利用大核注意力机制捕获水下图像的局部和全局信息,增强图像的语义表征,解决了水下图像的色彩失真问题;周露璐等^[11]为了加强模型不同通道之间的信息交互效率,提出了基于密集连接和残差学习的密集残差卷积神经网络,有效提高了水下图像的质量。

综上所述,在现有的水下图像增强方法中,概率模型的输入数据中存在噪声或伪影,可能会影响网络的效果,导致颜色补偿的准确性下降。另外,如果训练数据集的多样性和质量不高,模型的泛化能力可能会受到限制,导致在未知的水下场景中表现不理想。而基于深度学习的方法,采用卷积神经网络(convolutional neural

network, CNN)作为主要结构,导致模型的参数量大,训练和推理速度慢,不利于模型做边缘化设备部署。虽然网络考虑了输入图像的全局信息,但却忽略了如颜色、纹理、边缘与细节等关键特征,导致经过网络增强后的图像仍存在色彩恢复不理想,纹理不清晰和边缘出现伪影的问题。

因此,本文提出了基于关键特征传递的轻量化生成对抗网络水下图像增强模型,实现网络轻量化的同时,使增强的图像可以拥有理想的色彩效果,增强图像的清晰度并消除边缘伪影问题。本文提出了一种轻量级基于关键特征传递的水下图像增强模型(based on the generative adversarial network of key feature transfer, KFT-GAN),使用深度可分离卷积替代普通卷积,降低了模型的参数量,缓解了模型训练和推理速度慢的问题。为了让生成器网络能更好地关注图像的关键特征,设计了关键特征传递模块(key feature transfer module, KFT),使提取到的图像关键特征能更高效地在中间层传递,确保显著特征可以传递至后续层来优化模型性能,实现增强图像细节的目的。

1 网络模型

生成对抗网络(generative adversarial network, GAN)包含两个神经网络——生成器网络(generator, G)和判别器网络(discriminator, D),生成器以随机噪声作为输入,并生成合成数据,试图产生与真实数据相似的样本;判别器则被训练用来区分真实样本和生成的样本。生成器和判别器之间的对抗性相互作用促使两个网络不断学习进化^[12]。训练过程可以通过一个极小极大博弈来表示,其中生成器试图最小化判别器正确分类生成样本的概率,而判别器则旨在最大化其准确性。

本文根据图像处理的任务需求提出的轻量级 KFT-GAN 水下图像增强模型的生成器网络以水下退化图像作为输入,而判别器网络以生成器网络的生成图像 I_G 和参考图像 I_{GT} 作为输入。在模型训练阶段过程中,生成器与判别器相互监督,循环交替更新各自网络的参数。最终得到高效且稳定的生成器网络,并使用该生成器网络来增强恢复水下图像。KFT-GAN 模型整体结构如图 1 所示。

1.1 生成器网络

KFT-GAN 模型的生成器网络采用 U-Net^[13] 网络结构,包括编码器(encoder)和解码器(decoder)。编码器对输入图像进行特征提取,解码器通过反卷积操作进行图像重建,过程如图 2 所示。

KFT-GAN 的生成器网络设计为 5 层编解码结构。为了减少生成器网络的参数量和计算量,使网络更轻

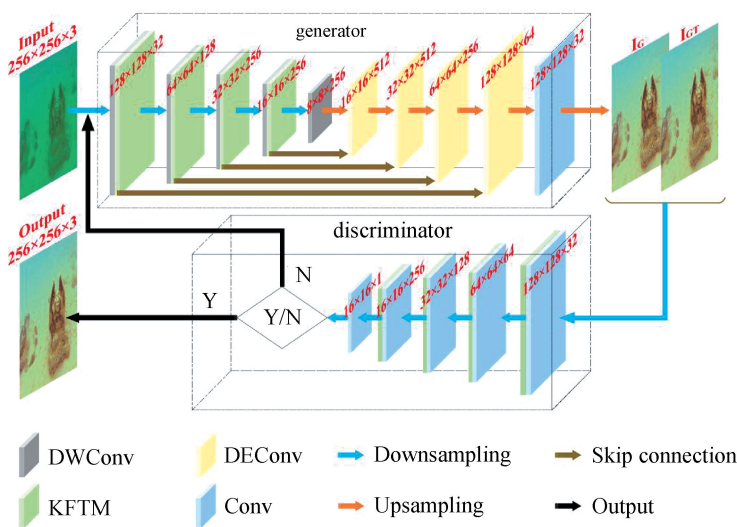


图 1 整体网络结构

Fig. 1 Network structure

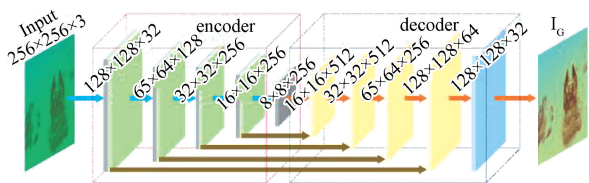


图 2 生成器网络结构

Fig. 2 Generator network structure

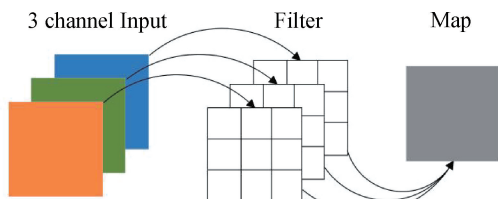


图 3 普通卷积操作

Fig. 3 General convolution

量化,在编码器网络使用深度可分离卷积 (depthwise separable convolution, DWConv)^[14]对输入特征图进行特征提取操作,并在 DWConv 后使用 KFT 模块对提取到的特征进行增强,使关键特征更好地传递至下一层,确保在图像在重构的过程中能表达更多的细节信息,提升重构图像的质量。编码器的输出直接作为解码器的输入,可以减少网络的信息丢失,且 KFT 模块获取的图像关键特征通过跳跃连接传递到上采样过程,增强上采样过程重建图像的细节表达,使网络可以更好地表示和强化图像细节信息,增强特征映射的表现力和鲁棒性。

1) 深度可分离卷积

在普通卷积操作中,如图 3 所示,当上一层具有 n 个通道时,进行卷积操作过程中每个滤波器 Filter 就需要包含 n 个卷积核 Kernel 与 n 个通道对应。每个滤波器进行一次卷积操作,事实上是多个卷积核与上一层对应通道的特征图进行卷积并求和,从而生成下一层的单个通道特征图。若在下一层中,需要生成 m 个通道的特征图,则需要 m 个滤波器与之对应。当输入特征图尺寸为 $H \times W \times 3$,卷积核尺寸为 3×3 ,要得到 x 个通道的输出特征图时,参数量 N 为 $N = x \times 3 \times 3 \times 3 = 27 \times x$ 。

深度可分离卷积,如图 4 所示,仅对普通卷积进行了一个微小的调整,但所需的参数量却显著减少了,有利于网络的轻量化设计。

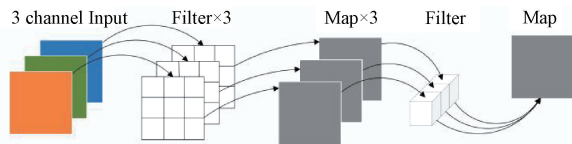


图 4 深度可分离卷积操作

Fig. 4 Depthwise separable convolution

深度可分离卷积分为两步,第一步为逐通道卷积,将上一层得到的多通道特征图,完全拆分为单通道特征图,再分别对每个单通道特征图进行卷积,然后将卷积得到的结果重新堆叠,这个过程只对上一层得到的多通道特征图的尺寸进行调整,不改变其通道数。第二步为逐点卷积,对逐通道卷积得到的特征图进行卷积,使用大小为 1×1 的卷积核,滤波器包含的 1×1 的卷积核数量与上一层特征图通道数一样。逐点卷积阶段单个滤波器得到单个通道特征图,要获得多个通道特征图,则要多个滤波器进行卷积操作。当输入特征尺寸为 $H \times W \times 3$,卷积核尺寸

为 3×3 , 要得到 x 个通道的输出特征图时, 参数量 $N = 3 \times 3 \times 3 + 1 \times 1 \times 3 \times x = 27 + 3 \times x$, 可以看出当特征图通道数越多时, 深度可分离卷积的参数量比普通卷积的参数量减少的越多, 模型更轻量化。

团队研究证明在设计网络结构中, 将普通卷积替换为深度可分离卷积, 减少了每层卷积运算所需的乘加运算量; 将标准卷积操作拆分为深度卷积和逐点卷积两个独立步骤, 能够显著降低计算复杂度和参数量, 进而降低了内存占用和计算需求, 不仅提高了模型的计算效率,

同时也减轻了过拟合的风险。因此, 本研究使用深度可分离卷积搭建轻量化的网络, 有助于提升推理速度, 保持较高的性能水平, 有利于高效的水下图像增强任务。

2) KFT 模块

为了解决生成器网络在捕捉图像边缘信息方面存在的局限性和无法有效捕捉全局上下文信息, 导致重构图像质量不佳的问题, 本文设计了一种关键特征传递模块 KFT, 结构如图 5 所示。

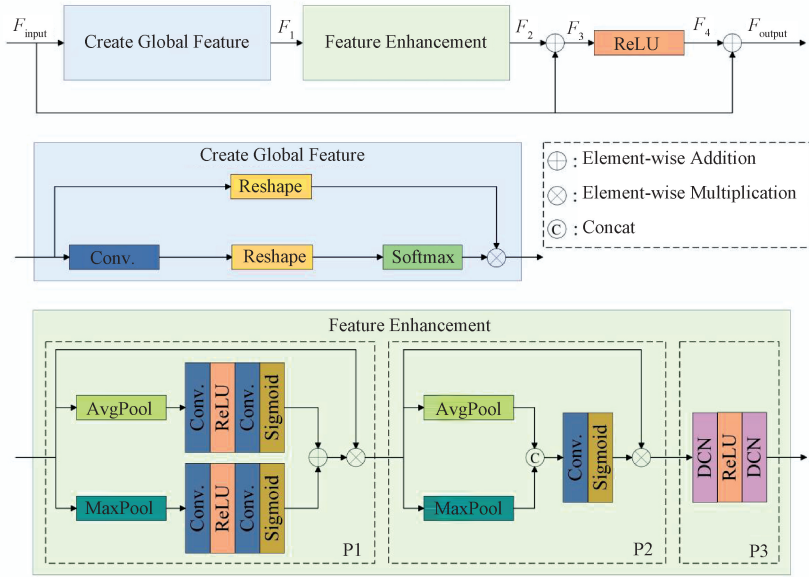


图 5 KFT 模块结构

Fig. 5 Structural diagram of the KFT module

该模块能有效地捕捉上下文信息并增强网络的特征表达能力, 通过选择性地抑制噪声及非关键特征的传播, 强化对重要信息的提取与保留, 同时有助于减轻无关或冗余信息对模型决策过程的潜在干扰, 确保输入特征中的显著特征可以传递至后续层优化模型性能, 从而实现更好的增强图像效果。KFT 模块整体流程如式 (1) 所示。

$$F_{\text{output}} = F_{\text{input}} \oplus \text{KFT}(F_{\text{input}}) \quad (1)$$

式中: F_{input} 代表输入特征; F_{output} 代表输出特征; KFT 为关键特征传递操作。

KFT 操作主要包括 3 个阶段:

(1) 全局特征创建阶段, 输入特征图 F_{input} 一方面直接通过 Reshape 操作得到重构特征, 另一方面经过一系列的 Conv、Reshape 和 Softmax 操作得到重构特征权重, 再通过逐元素相乘得到新的重构全局特征 F_1 。

(2) 特征强化阶段, 可分为步骤 P1、P2 和 P3。在步骤 P1 中对输入的特征图分别进行最大池化和平均池化操作, 压缩输入特征图的尺度, 以获取具有不同维度信息

的特征图, 接着将得到的特征图送入由 Conv、ReLU 和 Sigmoid 组成的序列操作中, 对输入特征图进行先降维再升维操作, 实现不同维度信息间的共享关联, 最后对特征图进行逐元素相加后与输入特征图进行逐元素相乘得到具有权重信息的输出特征图。计算过程如式 (2) 所示。

$$F_{\text{out}} = F_{\text{in}} \otimes (\text{Sigmoid}(\text{Conv.}(\text{ReLU}(\text{Conv.}(\text{AvgPool}(F_{\text{in}})))))) \oplus \text{Sigmoid}(\text{Conv.}(\text{ReLU}(\text{Conv.}(\text{MaxPool}(F_{\text{in}})))))) \quad (2)$$

式中: F_{in} 为输入特征图; F_{out} 为输出特征图。在步骤 P2 中, 先对步骤 P1 得到特征图分别进行最大池化和平均池化操作, 将得到的特征图进行通道维度堆叠后, 通过 Conv 调整通道数后使用 Sigmoid 激活函数进行激活, 再与步骤 P1 得到的输出特征图进行逐元素相乘得到输出特征图。计算过程如式 (3) 所示。

$$F_{\text{out}} = F_{\text{in}} \otimes (\text{Sigmoid}(\text{Conv.}(\text{AvgPool}(F_{\text{in}}) \text{MaxPool}(F_{\text{in}})))) \quad (3)$$

在步骤 P3 中, 步骤 P2 输出的特征图经过 Conv 和 ReLU 操作后得到增强后的输出特征图。

(3) 特征融合输出阶段, 阶段 (1) 和 (2) 获取的关键

特征 F2 与输入特征 F_{input} 进行逐元素相加后经过 ReLU 激活函数激活后得到保留的关键特征信息 F4, 最后与输入特征 F_{input} 进行逐元素相加后使关键特征输出传输到下一层。

KFT 模块的设计基于最大池化与平均池化在信息提取方面的互补性。其中, 最大池化能够有效突出显著特征, 抑制背景干扰, 有利于边缘信息和关键目标的保留; 而平均池化则倾向于保留全局的上下文信息, 帮助网络理解图像的整体结构与分布特征。通过融合这两种池化路径的特征, 实现了对局部显著信息与全局语义信息的协同建模, 从而提升了网络对重要特征的表达能力。

1.2 判别器网络

GAN 模型的判别器通过区分真实图像和生成图像来评估生成图像的真实性。它在评估生成输出的真实感方面起着至关重要的作用。通过对抗训练, 生成器旨在生成逼真的水下增强图像以欺骗判别器, 从而提高生成结果的整体质量。本文的判别器网络使用 PatchGAN^[15] 判别器网络, 并使用 KFT 模块对 PatchGAN 网络的判别器进行改进设计, 让判别器对输入图像的关键特征给予更多的关注, 增强判别器的判别能力, 改进后的判别器网络结构如图 6 所示。与传统的 GAN 判别器有所不同, 本文的判别器并不是对整个图像做出真假判断, 而是对输入图像的局部区域进行判断。这种局部判断方法能够让判别器更专注于图像的细节特征, 从而生成更加精细和真实的图像。

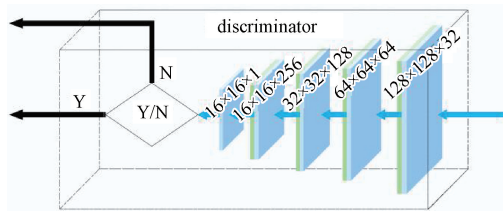


图 6 判别器网络结构

Fig. 6 Discriminator network structure

2 损失函数

为了使增强图像具有更好的视觉效果, 并保留更多的细节特征, 本文使用基于感知损失原理的混合损失函数, 包括感知损失、对抗损失和相似性损失 3 个部分, 定义如式(4)所示。

$$L = L_{VGG} + \min_c \max_D L_{WGAN}(G, D) + L_{L1} \quad (4)$$

式中: L_{WGAN} 为对抗损失函数^[16]; L_{L1} 为全局相似性损失函数^[17]; L_{VGG} 为感知损失函数^[18]; G 表示生成器网络; D 表示判别器网络。

2.1 L_{VGG} 感知损失

感知损失函数通过利用预训练的 VGG 网络来提取增强图像和参考图像之间的特征表示, 并通过计算这些特征之间的欧氏距离来评价增强图像的质量。感知损失函数能够捕捉到图像的高层次语义和结构信息, 确保生成的增强图像与参考图像之间具有相似的结构和纹理, 如式(5)所示。

$$L_{VGG} = \frac{\sum_{x=1}^W \sum_{y=1}^H (\alpha((I)_{x,y}) - \alpha((I^*)_{x,y}))^2}{HW} \quad (5)$$

式中: I 表示增强图像; I^* 表示参考图像; $\alpha((I)_{x,y})$ 和 $\alpha((I^*)_{x,y})$ 是由 VGG 网络提取的多尺度特征。

2.2 L_{WGAN} 对抗损失

GAN 损失函数的优化是基于散度进行的, 常导致网络损失的训练很不稳定, 带来梯度消失等问题, 导致模型训练崩溃^[19]。因此, 本文使用 WGAN^[20] 对抗损失函数, 该损失函数带有数梯度惩罚项, 可以使训练更稳定, 如式(6)所示。

$$L_{WGAN}(G, D) = E[D(y)] - E[D(G(x))] + \lambda_{GP} E[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \quad (6)$$

式中: \hat{x} 为生成图像与目标图像对应点之间直线上的样本; λ_{GP} 是梯度惩罚的权重, 设置的值为 10; x 表示整体网络的输入图像; y 表示整体网络的输出图像。

2.3 L_{L1} 全局相似性损失

全局相似性损失是训练模型的关键组成部分, 它量化了生成器网络生成图像 I_G 和相应的真实图像 I_{GT} 在像素上的不相似性。由于 L1 损失不容易引入模糊, 因此, 本文采用 L1 范数来量化生成的图像与其对应的真实图像之间的差值来优化判别器网络, 提高图像信息的一致性, 如式(7)所示。

$$L_{L1} = E[\|y - G(x)\|_1] \quad (7)$$

2.4 评价指标

为了衡量模型增强处理后的图像在颜色、结构、视觉细节和高级特征相似性的差异, 本文使用峰值信噪比、结构相似性指数、学习感知图像补丁相似性、局部边缘、视觉信息保真度作为评价指标。其中峰值信噪比 (peak signal to noise ratio, PSNR) 的值越高代表图像在降低噪声方面具有更好的优势; 结构相似性指数 (structural similarity index measure, SSIM)^[21] 测量增强图像与参考图像之间相似度, 值越高表示越好保护结构细节; 学习感知图像补丁相似性 (learning to perceive image patch similarity, LPIPS)^[22] 测量对象之间的感知相似性, 值越低表示网络越好感知相似性; 局部边缘 (lightness-order-error, LOE)^[23] 衡量的是增强图像和参考图像之间的局部边缘信息, 值越低越好; 视觉信息保真度 (visual

information fidelity, VIF)^[24] 衡量在增强图像中保存的视觉信息,更高的VIF值表明视觉细节保护更好。

3 实验与结果分析

本文实验是基于PyTorch框架的,所有实验均在NVIDIA RTX4090GPU上进行。训练批次大小设置为16,训练轮数为200 epoch。使用ADAM优化器,采用指数衰减学习策略,初始学习率为0.000 1。本文使用EUVP^[25]和UIEB^[26]数据集验证模型的有效性,其中EUVP数据集由20 000张水下图像数据组成,包括12 000张配对的图像数据和8 000张未配对的图像数据;UIEB数据集由950张水下图像数据组成,包括890张有配对参考图像的数据和60张没有配对参考图像的数据。再将本文提出的KFT-GAN模型和传统算法CLAHE^[27]、MSRCR^[28]、UDCP^[29]、深度学习算法

L²UWE^[30]、Shallow-UWNet^[6]、UWCNN^[31]、UWCNN-SD^[32]和FUuIEGAN^[26]进行对比,从主客观进行实验分析。

3.1 不同卷积的模型效率

本文使用DWConv代替普通卷积进行轻量化模型网络构建,为了探讨模型轻量化的效率和对模型性能的影响,对使用不同卷积方式的模型在EUVP数据集上进行了对比实验,结果如表1所示。从表1结果可以看出,使用DWConv对实现模型轻量化有显著效果,KFT-GAN(DWConv)模型的参数量比KFT-GAN(Conv)模型的参数量减少了60.6%,浮点运算量由51.42 GFLOPs降低到了24.58 GFLOPs。虽然SSIM值有略微降低,但是使用DWConv的模型整体性能表现要优于使用普通卷积的模型,PSNR值提高了0.414 9,推理时间快了0.09 s。说明使用DWConv的模型在模型性能不损失的情况下,能够显著降低模型的参数量和浮点运算量,提升计算效率。

表1 不同卷积方式的实验结果对比

Table 1 Comparison of different convolution

模型	参数量/(×10 ⁶)	运算量/GFLOPs	PSNR/dB ↑	SSIM ↑	处理时间/s
KFT-GAN(Conv)	8.02	51.42	20.969 3	0.744 2	0.13
KFT-GAN(DWConv)	3.16	24.58	21.384 2	0.741 3	0.04

3.2 损失函数性能分析

为了验证本文损失函数的有效性,在EUVP数据集上对不同损失函数进行了实验对比,如图7所示。

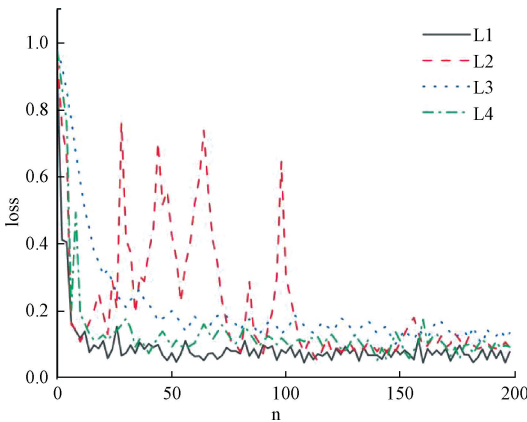


图7 不同损失函数训练过程对比

Fig. 7 Comparison of different loss function

其中L1是由对抗、感知和相似性组成的损失函数,L2是对抗损失函数,L3是由对抗损失和感知损失组成的函数,L4是由对抗和相似性组成的损失函数;loss为损失函数值,n为模型训练的迭代次数。从图7可以看出,L2曲线在训练过程中震动幅度大,且训练迭代次数超过100次后才慢慢进入稳定状态,说明只使用对抗损失函

数来训练并不能提供足够的信息来衡量模型的表现,导致模型训练不稳定。L3曲线和L4曲线相对L1曲线要平稳,主要原因是L3曲线在对抗损失的基础上加入了感知损失部分,使模型注重生成图像的高层次语义和视觉质量,训练过程中损失波动减少;L4曲线在对抗损失的基础上加入了相似性损失,通过衡量生成样本与真实图像在像素级别或高层次特征上的差异来进行优化,能够保证生成的图像在结构、细节等方面更接近目标图像,避免仅通过对抗损失生成模糊或失真的图像。从L1曲线看出,在对抗损失函数的基础上加入感知损失函数和相似性损失函数使模型的训练过程变得更加平稳,曲线的整体波动幅度较小,且损失函数值更低。

由表2可知,本文使用的损失函数的PSNR值是21.384 2,SSIM值是0.741 3,明显优于另外3种损失函数的值,说明在对抗损失的基础上引入感知损失和相似性损失后,生成图像在细节上更加清晰,且能够更好地保留图像的原始风格或结构,使得生成图像更符合实际情况。

表2 不同损失函数实验结果对比

Table 2 Comparison of different loss function

损失函数	PSNR/dB ↑	SSIM ↑	LOE ↓	LPIPS ↓	VIF ↑
L4	19.459 3	0.689 5	241.376	0.359 8	0.386 5
L3	19.085 4	0.662 4	263.605	0.321 7	0.402 7
L2	18.937 2	0.621 9	285.712	0.374 2	0.354 1
L1	21.384 2	0.741 3	212.833	0.301 6	0.437 5

3.3 主观评价

各对比模型在 EUVP 数据集上的增强效果如图 8 所示,从图 8(b)~(d)可以看出,传统方法对图像进行增强的时容易出现过渡增强的问题,导致增强后的图像加重了色偏问题,与图 8(k)对比可以看出,传统方法在恢复图像色彩方面的能力较差。从图 8(e)~(i)可以看出,深度学习方法在恢复图像的色彩方面具有较好的方向

感,但仍然存在色偏的问题,且增强的图像在清晰度方面欠佳。从图 8(j)可以看出,本文提出的 KFT-GAN 模型增强的图像具有更好的视觉效果,与其他模型相比,增强图像的色彩更接近真实环境的色彩,与图 8(k)参考图像相比具有更好的清晰度。本文模型增强的图像无论是图像物体的轮廓边缘还是纹理部分,都是对比模型中最佳的。

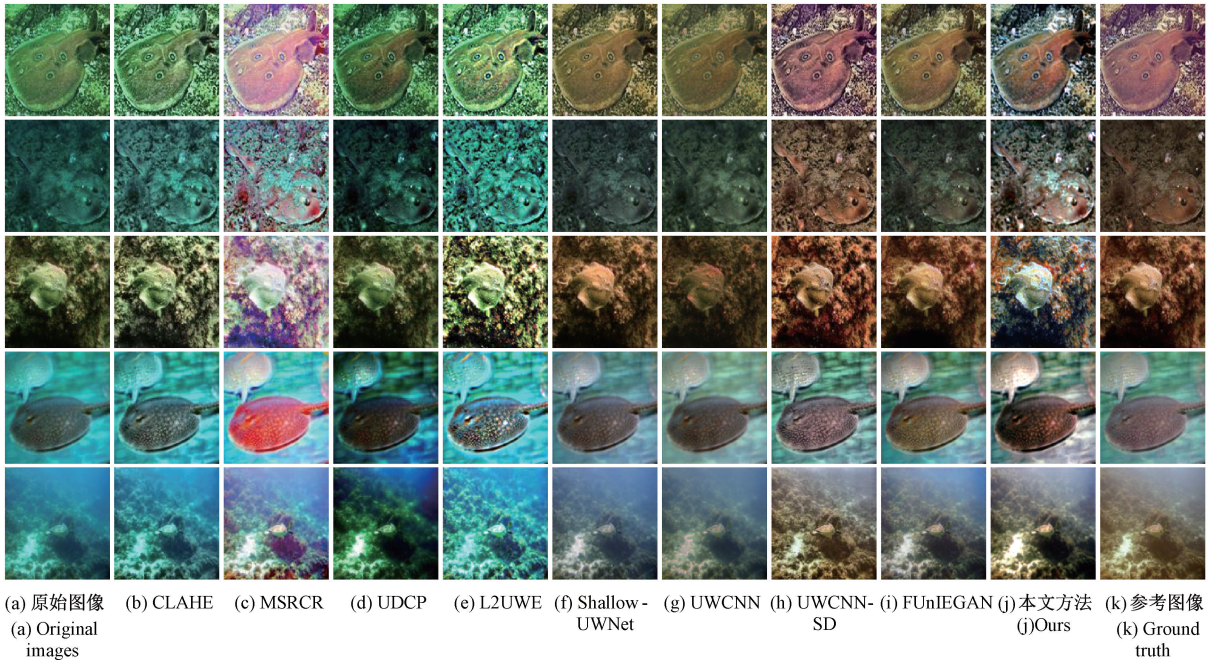


图 8 EUVP 数据集效果对比

Fig. 8 Comparison chart of enhancement results on the EUVP dataset

各对比模型在 UIEB 数据集上的增强效果如图 9 所示。从图 9 可以看出本文模型增强处理后的图像在颜色上更接近真实自然的色彩。与其他模型处理的图像相比,本文模型处理的图像具有更好的清晰度和对比度,噪声的抑制效果明显优于其他模型。

3.4 客观评价

由表 3 和 4 的定量比较实验结果可以看出,与其他模型相比,本文提出的 KFT-GAN 模型在 EUVP 数据集和 UIEB 数据集上的 PSNR、SSIM、LPIPS、LOE、VIF 指标均取得最优的结果,表明本文提出的方法在有效地保留边缘细节和纹理结构视觉信息的同时,在抑制伪影和恢复颜色信息方面都要优于其他模型。实验结果表明了本文提出的基于关键特征传递的生成对抗网络模型在水下退化图像增强中的鲁棒性和有效性。

实验表明,本文提出的 KFT-GAN 在 EUVP 数据集和 UIEB 数据集上表现出更好、更稳定的性能,能准确地表

达目标物体和背景颜色,有效提升细节清晰度。所以本文提出的模型对水下退化图像的增强效果更好,具有较好的工程适应性。

表 3 EUVP 数据集定量比较

Table 3 Quantitative comparison of EUVP dataset

Method	PSNR/dB ↑	SSIM ↑	LOE ↓	LPIPS ↓	VIF ↑
CLAHE	15.213 6	0.521 2	310.468	0.498 4	0.299 1
MSRRCR	16.381 2	0.675 8	237.960	0.339 2	0.341 8
UDCP	15.738 5	0.619 4	240.812	0.321 5	0.326 7
L ² UWE	13.881 7	0.398 1	398.637	0.639 1	0.353 8
Shallow-UWNet	13.979 6	0.401 7	319.723	0.5240	0.339 1
UWCNN	14.395 1	0.413 1	413.191	0.532 9	0.268 2
UWCNN-SD	16.907 9	0.632 5	236.374	0.349 3	0.329 4
FUnIEGAN	18.218 4	0.597 4	289.536	0.409 7	0.313 9
KFT-GAN	21.384 2	0.7413	212.833	0.301 6	0.437 5

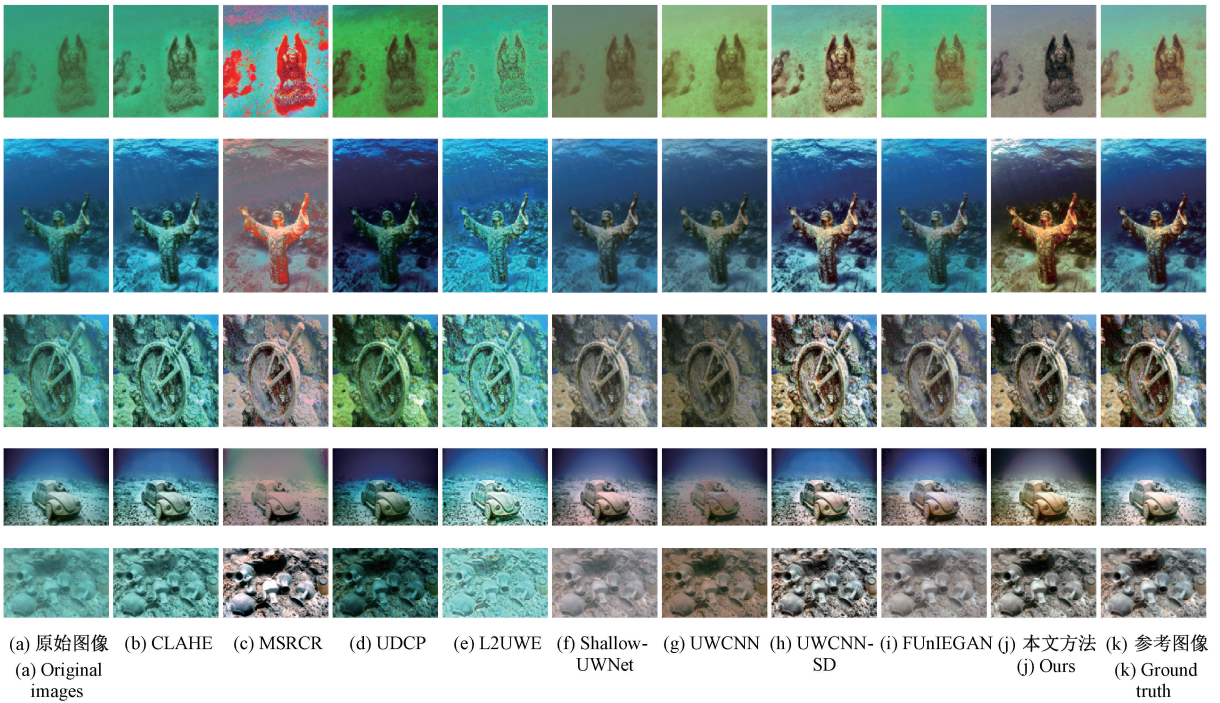


图 9 EUVP 数据集效果对比

Fig. 9 Comparison chart of enhancement results on the EUVP dataset

表 4 UIEB 数据集定量比较

Table 4 Quantitative comparison of UIEB dataset

Method	PSNR/dB ↑	SSIM ↑	LOE ↓	LPIPS ↓	VIF ↑
CLAHE	14.698 3	0.472 5	338.732	0.549 7	0.271 4
MSRCR	15.997 6	0.639 5	254.695	0.360 3	0.311 2
UDCP	15.197 9	0.589 6	248.139	0.359 8	0.294 1
L ² UWE	13.298 7	0.359 7	421.196	0.699 2	0.316 5
Shallow-UWNet	13.493 5	0.360 2	348.989	0.594 9	0.304 8
UWCNN	12.201 4	0.349 1	446.372	0.580 4	0.230 3
UWCNN-SD	15.099 8	0.595 3	248.466	0.370 8	0.294 6
FU_nIEGAN	16.615 2	0.553 4	306.897	0.463 1	0.257 4
KFT-GAN	18.025 6	0.688 9	244.591	0.348 5	0.317 9

4 结 论

针对 CNN 网络模型参数量大且容易忽略图像关键特征的问题,本文设计了 KFT 模块,并结合 DWConv 构建了轻量化的基于 KFT 模块的生成对抗网络水下退化图像增强模型 KFT-GAN。通过网络间的下采样和跳跃连接实现了关键特征的高效传递,提高了解码阶段图像重建的质量。此外,得益于基于感知损失理论的混合损失函数来引导图像重建,增强了重建图像的质量,视觉上具有更好的色彩和细节恢复效果。

本文从主观和客观两个方面验证了提出模型在水下

图像增强方面具有最优的视觉效果和最好评价指标。从研究中发现还存在以下问题:1) 对抗训练的稳定性,下一步将继续研究更稳定的对抗训练策略,如谱归一化等技术;2) 跨域适应与泛化能力,当前模型可能在水下特定场景下表现良好,但在不同水域或不同光照条件下的泛化能力可能有限。为此可以研究跨域适应技术,减少对参考图像的依赖性,使该模型具有更好的泛化能力,更好地适应不同的水下环境。

在未来的研究中,将侧重于水下图像特性的优化物理模型,结合水下物理光学先验知识来改进图像增强算法,进一步优化模型的实时性与计算效率,研究如何在保证图像质量的前提下,使其能够在嵌入式设备或移动设备上实时运行。

参考文献

[1] ZHANG H, LI X C, YAO G P, et al. A pulse delay control method by external trigger to underwater illumination [J]. Journal of Nanoelectronics and Optoelectronics, 2022, 17(8): 1185-1194.

[2] FU C, YUAN H, XU H J, et al. TMSO-Net: Texture adaptive multi-scale observation for light field image depth estimation [J]. Journal of Visual Communication and Image Representation, 2023, 90: 103731.

[3] HU K, WENG C H, ZHANG Y W, et al. An overview of underwater vision enhancement: From traditional methods to recent deep learning [J]. Journal of Marine

- Science and Engineering, 2022, 10(2): 241.
- [4] RAO Y, LIU W, LI K, et al. Deep color compensation for generalized underwater image enhancement[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 34(4): 2577-2590.
- [5] CHANG B C, LI J J, WANG H Y, et al. Msap: Multi-scale attention probabilistic network for underwater image enhancement network [J]. Signal, Image and Video Processing, 2024, 18(Suppl 1): 653-661.
- [6] NAIK A, SWARNAKAR A, MITTAL K. Shallow-wnet: Compressed model for underwater image enhancement (student abstract)[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(18): 15853-15854.
- [7] ALYASSERI Z A A, GHALIB R, JAMIL N, et al. Lighting enhancement of underwater image using coronavirus herd immunity optimizer [J]. Alexandria Engineering Journal, 2024, 91: 115-125.
- [8] YANG A P, WANG CH C, WANG J B, et al. Zero-reference single underwater image enhancement [J]. Multimedia Tools and Applications, 2023, 82 (30): 46423-46438.
- [9] CHEN Y H, LI Q F, LU D X, et al. A novel underwater image enhancement using optimal composite backbone network[J]. Biomimetics, 2023, 8(3): 275.
- [10] XU S H, WANG J H, HE N, et al. Underwater image enhancement method based on a cross attention mechanism[J]. Multimedia Systems, 2024, 30(1): 26.
- [11] 周露珊,赵磊,李恒,等. 基于轻量级密集残差网络的水下图像增强 [J]. 电子测量与仪器学报, 2023, 37(1):70-77.
- ZHOU L S, ZHAO L, LI H, et al. Underwater image enhancement based on lightweight dense residual network[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(1): 70-77.
- [12] SHI W Z, TAO F, WEN Y. Structure-aware deep networks and pixel-level generative adversarial training for single image super-resolution[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-14.
- [13] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]. International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer international publishing, 2015: 234-241.
- [14] HOWARD A G, ZHU M L, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J]. arXiv preprint arXiv: 1704.04861, 2017.
- [15] NGUYEN T, YOO M. Patchgan-based depth completion in autonomous vehicle [C]. 2022 International Conference on Information Networking (ICOIN). IEEE, 2022: 498-501.
- [16] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks [C]. International Conference on Machine Learning. PMLR, 2017: 214-223.
- [17] 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.
- [18] WANG Y, FAN H, LIU S, et al. BA-GAN: Block attention GAN model for Underwater Image Enhancement[C]. 2021 IEEE International Conference on Unmanned Systems (ICUS). IEEE, 2021: 809-813.
- [19] YU X L, QU Y Y, HONG M. Underwater-GAN: Underwater image restoration via conditional generative adversarial network [C]. International Conference on Pattern Recognition. Cham: Springer International Publishing, 2018: 66-75.
- [20] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein gans[J]. Advances in Neural Information Processing Systems, 2017, 30.
- [21] 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.
- [22] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 586-595.
- [23] SHEIKH H R, BOVIK A C. Image information and visual quality [J]. IEEE Transactions on Image Processing, 2006, 15(2): 430 - 444.
- [24] WANG S, ZHENG J, HU H M, et al. Naturalness preserved enhancement algorithm for non-uniform illumination images [J]. IEEE Transactions on Image Processing, 2013, 22(9): 3538-3548.
- [25] 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.
- [26] 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.

- [27] 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.
- [28] ZHANG S, ZENG P, LUO X, et al. Multi-scale retinex with color restoration and detail compensation [J]. Journal of Xi'an Jiaotong University, 2012, 46(4): 32-37.
- [29] DREWS P L J, NASCIMENTO E R, BOTELHO S S C, et al. Underwater depth estimation and image restoration based on single images [J]. IEEE Computer Graphics and Applications, 2016, 36(2): 24-35.
- [30] MARQUES T P, ALBU A B. L2uwe: A framework for the efficient enhancement of low-light underwater images using local contrast and multi-scale fusion [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 538-539.
- [31] LI C, ANWAR S, PORIKLI F. Underwater scene prior inspired deep underwater image and video enhancement[J]. Pattern Recognition, 2020, 98: 107038.
- [32] WU S, LUO T, JIANG G, et al. A two-stage underwater enhancement network based on structure decomposition

and characteristics of underwater imaging [J]. IEEE Journal of Oceanic Engineering, 2021, 46(4): 1213-1227.

作者简介



麦仁贵, 2018年毕业于广东技术师范学院, 现为广东海洋大学硕士研究生, 主要研究方向为深度学习与图像处理。

E-mail: renguim@163.com

Mai Rengui received his B. Sc. degree from Guangdong Polytechnic Normal University in 2018. Now he is a M. Sc. candidate at Guangdong Ocean University. His main research interests include deep learning and image processing.



王骥(通信作者), 2010年毕业于广东工业大学, 现为广东海洋大学教授, 主要研究方向为海洋无线传感器网络(WSN)、人工智能与信息处理。

E-mail: 13902576499@163.com

Wang Ji (Corresponding author) received his B. Sc. degree from Guangdong University of Technology in 2010. Now he is a professor at Guangdong Ocean University. His main research interests include marine wireless sensor networks (WSN), artificial intelligence and information processing.