

多尺度注意力机制的水下图像增强算法

王士川, 杜 玲

天津工业大学计算机科学与技术学院, 天津

收稿日期: 2021年8月12日; 录用日期: 2021年10月7日; 发布日期: 2021年10月14日

摘 要

由于不同波长的光在水环境中的衰减差异, 原始的水下图像会出现偏绿、偏蓝和白雾的问题。针对水下图像增强出现的上述问题, 本文通过结合多尺度密集块和注意力机制, 提出了一个生成式对抗网络。在生成器中使用多尺度残差密集块, 在后端使用注意力机制块来关注前端的特征, 前者用来提高网络学习能力, 后者增强网络的特征选择能力。判别器的每一层都使用谱归一化, 放宽判别标准, 以减少异常数据引起的梯度崩溃。同时, 结合多种损失函数来增强生成对抗网络的泛化能力。提出的算法在公共数据集上进行了测试, 并在实际应用中得到了验证。实验结果表明, 该算法比现有算法有更好的性能。

关键词

水下图像增强, 生成对抗网络, 深度学习

Underwater Image Enhancement Based on Multi-Scale Attention Mechanism

Shichuan Wang, Ling Du

School of Computer Science and Technology, Tiangong University, Tianjin

Received: Aug. 12th, 2021; accepted: Oct. 7th, 2021; published: Oct. 14th, 2021

Abstract

Due to the attenuation difference of different wavelengths of light in the water environment, the original underwater images show greenish, bluish and white haze problems. For underwater im-

age enhancement, this paper proposes a generative adversarial network by combining multi-scale dense blocks and attention mechanism. The residual multiscale dense block is used in the generator, and the attention mechanism block is used in the back-end to focus on the front-end features, which can improve the learning ability and enhance the network's ability to select features, respectively. Each layer of the discriminator uses spectral normalization and relaxes the discriminant criteria to reduce the gradient collapse caused by anomalous data. At the same time, multiple loss functions are combined to enhance the generalization ability of GAN networks. The proposed algorithm is tested on public datasets and validated in practical applications. Experimental results show better performance than existing algorithms.

Keywords

Underwater Image Enhancement, Generative Adversarial Network, Deep Learning

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来, 水下图像增强已经吸引了研究人员的广泛关注。水下图像的质量受光的影响最大, 不同波长的光传播的距离不同。水深越深, 长波在水中的衰减就越严重, 影响也就越强[1]。此外, 图像的衰减实际上是多种因素综合作用的结果, 如水温、微生物的种类和数量。由于水下图像既是水下信息的载体, 又是水下场景的表现形式, 在探测和感知水下环境方面起着不可替代的作用。因此, 水下图像增强仍然是一项亟需解决的艰巨任务。

深度学习在视觉任务中的成功, 如目标检测和图像分割, 激发了越来越多的深度模型在图像恢复中的应用[2]。然而, 对于水下任务来说, 数据集是一个难以解决的问题。现有的水下数据集相对稀少, 而且参考图像的质量很难与自然图像相比。生成对抗网络(GAN)比其他基于深度学习的模型显示出更好的视觉效果。它将图像恢复问题转化为图像生成问题, 其独特的自主生成特性, 缓解了高质量数据集的一些压力。目前, 已经提出许多有效的方法用于水下图像增强任务。

然而, 对于基于 GAN 的水下图像增强方法, 仍然存在一些限制。首先, 用于训练的水下图像数据集的多样性是不够的。根据[3]中测量的十个不同的水体的结果, 淡水和咸水水体之间存在很大的差异。不同纬度的水体和不同强度的环境光导致同一水下目标的颜色有所不同。其次, 生成的模型对特殊的数据特征没有给予足够的重视。根据经典的 Jaffe-McGlamery 水下成像模型[4], 白色雾状模糊团往往出现在浅水区。随着水深的增加, 不同波长的光的衰减是非线性的, 色彩对比度逐渐降低。这导致了在一定深度下的显著内容损失, 使得 GAN 模型很难将白雾从图像场景中区分出来, 增强结果会出现对比度偏差并出现难以解决的模糊感。第三, 由于现有的 GAN 模型采用动态平衡作为收敛的判断结果, 因此结果不够稳定, 传统的对抗性损失在训练的早期阶段很难收敛, 这会对增强效果产生影响。

在本文中, 我们提出了一个条件生成对抗网络(c-GAN), 它结合了多尺度密集块和注意力机制, 主要贡献如下:

- 1) 提出了一个结合多尺度密集块和注意力机制块的生成器网络, 直接从退化的图像中获得增强的结果, 而不需要预先构建水下退化模型。
- 2) 使用多个内容损失函数来约束网络, 限制 GAN 网络的梯度。放宽判别标准, 使网络更加关注关

键特征, 优先优化主要内容, 更稳定、快速地达到收敛状态。

3) 采用两个与人类感官高度一致的非参考评价指标对实验结果给出了主观和客观评价, 证明了所提出的算法在水下图像增强中的优越性。

2. 相关工作

水下图像增强是一项艰巨的任务, 主要由于水下复杂的光照环境会严重降低可见度并且导致色彩失真。近年来提出了许多水下图像恢复和增强的方法, 现有的方法可归纳为以下两类。

2.1. 基于传统的方法

传统方法主要使用基于物理散射模型的人工方法。Wei 等[5]使用水下光的衰减先验模型来估计水下场景图像的深度信息, 从而可以获得预测的背景光和透射图信息。Huang 等[6]将图像恢复分为两部分: 对比度恢复与颜色调整。首先在 RGB 颜色空间中使用动态参数均衡 G 和 B 通道, 然后在 CIE-Lab 色彩空间中调整各个分量来恢复颜色。与[5]和[6]不同, Ancuti 等[7]尝试以多尺度融合的方式统一强度校正和边缘锐化。基于先验的人工方法缺乏概括性, 如果不满足该假设, 则无法获得良好的性能。因此, 在野外, 它们通常表现较差, 例如不良的伪影、颜色失真等。

2.2. 基于深度学习的方法

深度学习技术所拥有的强大特征学习能力是有目共睹的, 其主要通过学习大量数据特征拟合不同图像信息的映射关系。根据使用的不同网络框架, 将其大致划分为基于 CNN 的方法和基于 GAN 的方法。

Li 等人[8]对比了多种水下图像增强算法的主观结果, 构建了一个真实世界水下图像数据集, 并且提出了一种基于 CNN 的融合增强算法, 通过对原始图像进行直方图均衡化、伽马矫正、自适应白平衡等操作得到多种增强图像, 并训练学习不同权重从而得到最终的融合结果。Hou 等人[9]提出了一种在图像的传输图像域使用残差学习的水下图像增强网络, 将水下图像增强任务建模为同时学习传输图和场景残差, 该方法可以将先验知识和数据信息整合在一起以研究水下图像的潜在分布。Islam 等[10]提出了具有端到端可训练模型的一种统一方法 Deep-SESR, 将水下图像增强和单图像超分辨率结合起来, 最高可以得到 4 倍分辨率的增强图像。该体系结构结合了密集残差子网和辅助注意力网络, 以促进网络学习局部密集的信息特征。

在[11]中, Uplavikar 等人提出了一种统一了十种水类型的 GAN 模型, 该模型通过一个干扰网络对编码器的输出进行判别, 迫使编码器对不同域的水下图像的输出分布更接近, 从而消除不同类别对解码器的影响。Li 等[12]从图像风格转换收到启发, 改进了 CycleGAN 网络(WCCT), 添加一项 SSIM 损失作为内容损失, 联合 CycleGAN 的循环一致性损失和对抗损失, 无需成对的数据集即可将水下图像的内容和结构接近空气中的图像。Fabbri 等人[13]直接使用 CycleGAN, 通过对 ImageNet 的子集[14]生成成对图像以解决训练集的稀少问题, 使用 WGAN 损失约束 GAN 网络, 梯度损失和 L1 损失约束内容, 训练 UGAN 恢复网络。

不同于上述方法, 本文使用水下失真图像和对应参考图像训练模型, 使用了多尺度密集块和通道-空间注意力机制来提升质量, 只需一个训练模型就可以在不同类型的水下图像中取得良好的效果。

3. 多尺度注意力水下图像增强算法

注意力模型在视觉任务、机器翻译和统计学习等深度学习领域广泛使用, 其对有序和无序数据均表现出较强效果。对于输入图像, 注意力机制会对输入特征计算权重, 并根据此权重对通道和空间位置进行筛选, 通过此种筛选方式, 可以使网络重点关注更加有用的特征, 本质上是使得网络自主选择更具效

果的卷积核, 其在编码器-解码器结构中应用广泛。通过引入多种注意力机制分支, 可以筛选出更为重要的特征信息。本文提出的网络通过利用 GAN 网络端到端的生成机制, 直接将失真的水下图像转换为未失真的增强结果, 整体框架如图 1 所示。x 为输入的失真水下图像, y 为与之成对的清晰的参考图像, G 为生成器模型, D 为判别器模型, 将 x 输入到生成器中得到生成器生成的图像 $G(x)$, 将生成图像输入判别模型 D 中, 得到判别结果概率, 将判断结果以反馈的形式传到生成模型 G 中指导生成模型训练。

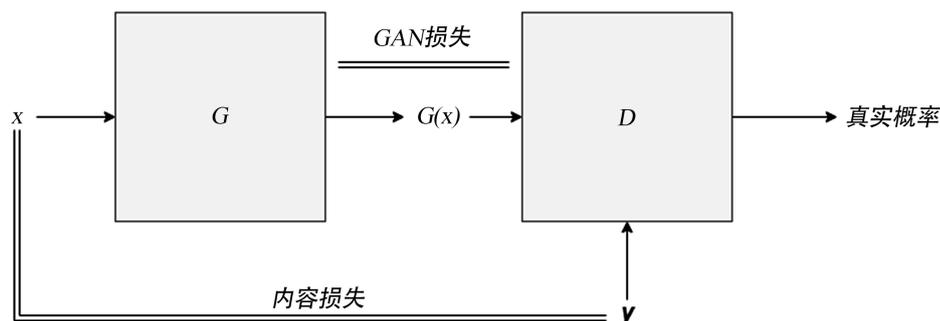


Figure 1. Overall framework
图 1. 整体框架

本文提出的方案使用对抗损失、L1 损失与 L2 损失组合的像素级别的内容损失、SSIM 损失和特征损失组合的结构损失共同限制整个网络来生成视觉上质量良好的图像。

3.1. 网络结构

3.1.1. 生成器网络

本文网络基于 U-Net [15] 结构, 但减少了重复块的数量, 压缩了特征图的大小, 如图 2 所示。网络中部的两个多尺度密集块(MSDB)将生成器网络划分为编码器和解码器两个部分。编码器阶段, 特征图每经过一次卷积层, 会使得其特征数加倍并且特征尺寸减半。在编码器末端, 特征图尺寸降低至 $16 \times 16 \times 256$ 。与 U-Net 相比, 它被减少到原来的四分之一, 从而简化了计算工作。整个网络省略了池化层, 改用步长为 2 的卷积层代替。事实证明, 使用卷积层代替池化层, 可以让卷积核在一张更大的输入图片上滑动[16], 得到每个区域的输出, 突破输入尺寸的限制, 并且池化层的使用会直接导致细节特征的丢失。在解码器阶段, 特征图每进行一次尺寸加倍, 会使特征图通道数减半。与编码器对应, 在解码器阶段使用上采样代替了转置卷积层, 以平衡编码器和解码器的运算。在编码器和解码器的对应位置, 全部使用跳跃连接。这消除了深度网络中的梯度消失问题, 使解码器能够获得最直接的特征信息。

广泛使用的感知结构[17]将不同比例的特征在块的结尾处简单地串联在一起, 这导致了部分特征图的利用不足。此外, 文献[18]提出了一个深度残差网络结构的框架, 可以轻松享受更多的特征结果。文献[19]提出了一种应用于水下图像的密集块, 对图像进行多尺度特征提取和特征复用。受上述结构的启发, 本文提出了一种改进的多尺度密集块结构, 将其嵌入编码器和解码器之间, 使用两个多尺度密集块作为最终结果。图 3 描述了本文的多尺度密集块的详细结构。如图 3 所示, 在该区块的前两层中, 分别使用两个具有内核和内核的卷积层, 每个卷积层后都跟 LeakyReLU 激活层, 使用不同内核可以提取更多的特征。本文尝试了内核数从 64 到 256 的多种组合, 最终选择 64 作为卷积核个数, 这有利于特征的高效传输。然后, 将两层的结果与原始输入进行串联。这样, 网络只需要提取输入和差异特征之间的差异特征。每个串联过程由包含原始输入的多尺度特征图组成, 残差特征可以得到充分的利用, 这有助于注意力网络对特征的过滤。为了便于串联操作, 多尺度密集块中的每一层都使用步长为 1 的卷积核, 块末尾的卷

积层将特征图的大小调整为与输入的大小一致, 这使得该块可以自由堆叠, 便于调整和移植。多尺度密集块的外部添加了跳跃连接, 以进一步鼓励信息流动。作为一个可堆叠的块状结构, 经过对不同数量的块状结构进行了尝试, 最终, 本文选择用两个块状结构线性连接来增强网络。

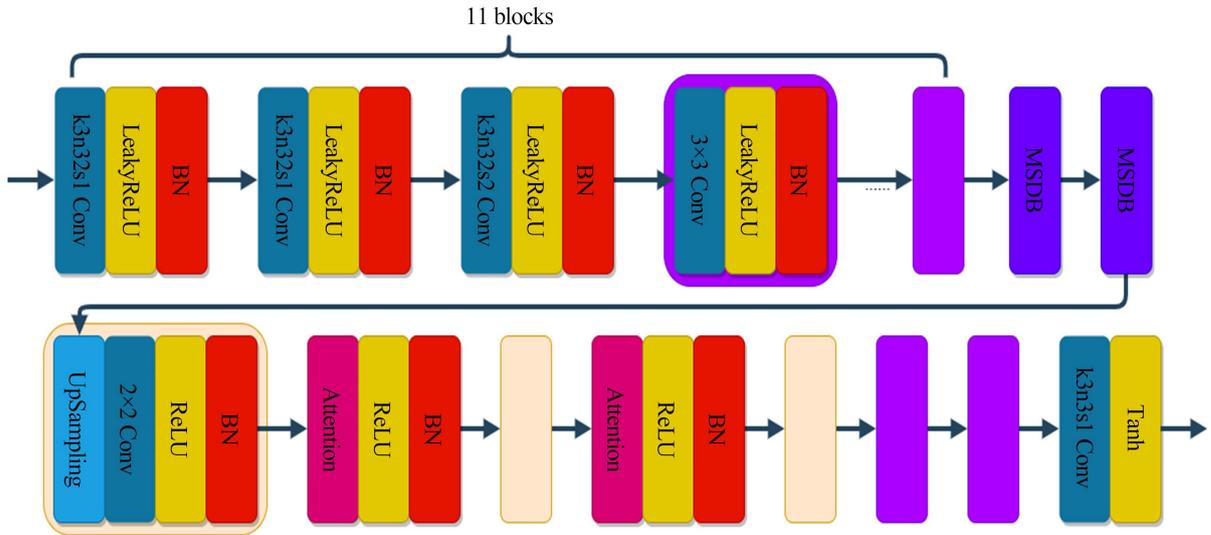


Figure 2. Generator network structure
图 2. 生成器网络结构

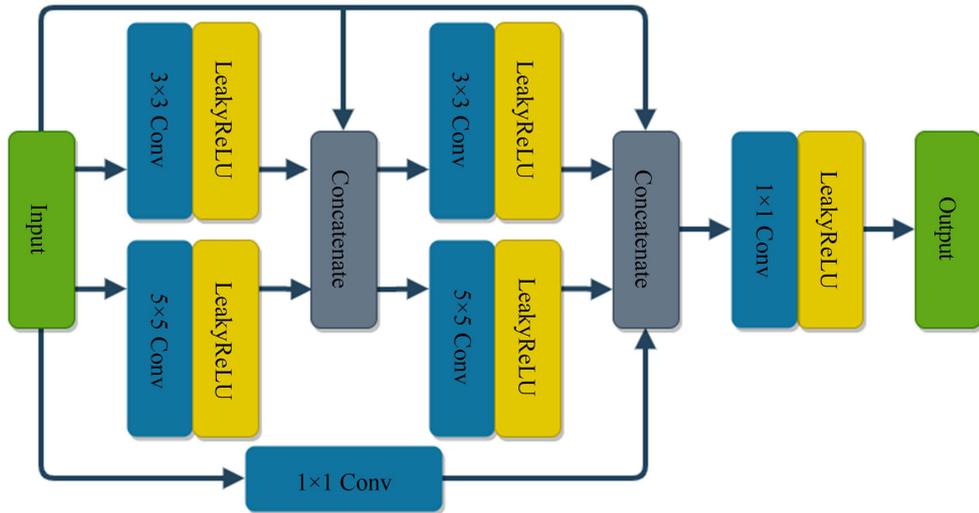


Figure 3. Multi-scale dense block structure
图 3. 多尺度密集块结构

在编码器第 8 和第 10 个块和解码器第 2 和第 4 个块的位置, 特征图尺寸各自为 64 和 128, 本文使用两种注意力分支进行并联处理。因此, 特征图可以沿着通道和空间进行重新校准。该部分的结构如图 4 所示。通道注意分支的作用是重新分配每个通道的权重, 突出具有重要特征的通道, 削弱具有不重要特征的通道。空间注意力分支的作用是强调图像的主要内容区域, 忽略图像的次要区域。前端和后端特征可以分别表示为 $F_f = [f_{f_1}, \dots, f_{f_i}, \dots, f_{f_c}]$ 和 $F_b = [f_{b_1}, \dots, f_{b_i}, \dots, f_{b_c}]$ 。其中 C 是处于前后两端的输出通道数, 前端特征和后端特征拥有相同的特征图尺寸。

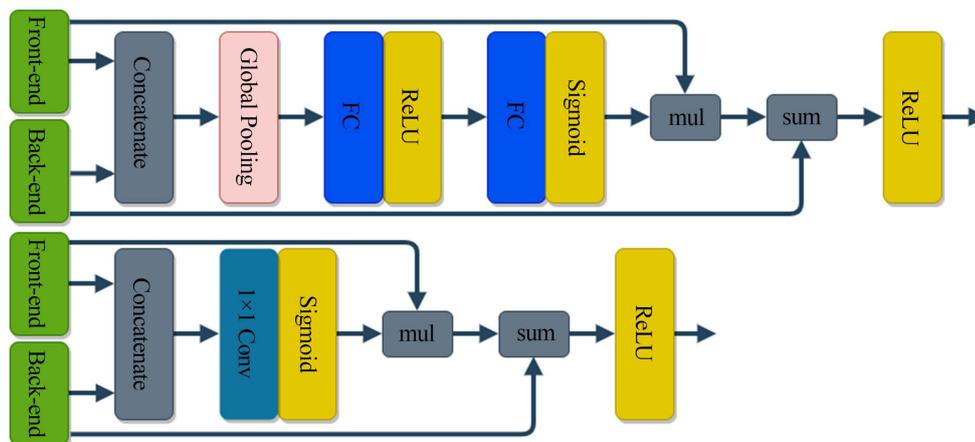


Figure 4. Channel attention branch (top) and spatial attention branch (bottom) structure
图 4. 通道注意力分支(上)和空间注意力分支(下)结构

通道注意力分支的作用是为前端和后端特征提取一组权重, 并将其作用于后端通道。首先, 将相应的特征图 F_b 和 F_f 串联以保留所有的特征信息。然后, 该张量依次通过全局平均池化层和两个全连接层, 经 Sigmoid 层将每个通道的权重范围 $W_{ch} = [w_1, w_2, \dots, w_C]$ 限制为 $[0, 1]$, 因此得到的特征图为 $1 \times 1 \times C$ 张量权重。最后, 将该权重与 F_f 的每个通道的所有元素相乘, 得到 F_f 经过权重调整后的结果, 该结果与 F_b 矩阵相加, 并经过 ReLU 层激活, 即可得到通道融合特征分支的结果 $F_{channel} = \text{ReLU}(F_b + W_{ch} * F_f)$ 。

空间特征融合分支的作用是将特征图的所有通道压缩成单通道特征, 调整后端特征所有通道中处于不同位置的内容权重。首先, 两组特征图 F_f 和 F_b 在通道维度上被连接。接下来, 通过 1×1 卷积, 将特征图压缩成 2 维的平面特征图像。然后, 经过 Sigmoid 层, 将 2 维特征的像素权重范围 $W_{sp} = [w_s^{1,1}, \dots, w_s^{i,j}, \dots, w_s^{H,W}]$ 限制为 $[0, 1]$ 。将此结果与 F_f 的每个通道逐一相乘, 得到加权后的结果。然后, 该结果与 F_b 矩阵相加, 并经过 ReLU 层激活, 即可得到空间融合特征分支的结果 $F_{spatial} = \text{ReLU}(F_b + W_{sp} * F_f)$ 。

最后, 本文通过逐元素相加 $F_{channel}$ 和 $F_{spatial}$, $F_{cs-fusion} = F_{channel} + F_{spatial}$ 来获得融合的特征图 $F_{cs-fusion}$ 。这将多尺度密集块所得到的充足后端特征进行重新调整, 使网络更加集中于有意义的特征。通过自主获得权重调整后的特征图, 可以加速网络的收敛, 获得良好的内容。

3.1.2. 判别器网络

所提出的判别器网络结构受到马尔可夫判别器(PatchGAN) [20]的启发。网络完全由卷积层组成, 可以对任意大小的图像进行判别。前两个卷积层的步长为 1, 其余卷积层的步长为 2, 每经过一次卷积, 特征图数量增加一倍。所有的卷积核尺寸都设置为偶数, 便于更好地提取边缘特征。每个卷积层都使用谱归一化[21]进行梯度约束, 以限制判别器的 Lipschitz 常数。文献[22]显示, 具有谱归一化的判别器比没有谱归一化的判别器具有更好的客观质量分数。卷积层后紧跟 LeakyReLU 激活层, 用以稳定上层的输出。在上述网络设置下, GAN 网络具有非常强的稳定性。

3.2. 损失函数

GAN 的生成器网络生成一个未失真的图像, 用来欺骗判别器。判别器接收一组未知图像和原始失真图像串联的图像数据, 并对这组图像进行判别。GAN 的训练过程是交替优化生成网络和判别器网络。对抗网络的损失函数如公式(1)所示:

$$\mathcal{L}_{\text{GAN}} = \mathbb{E}_{x \sim P_{\text{train}}(x)} [\log D(x)] + \mathbb{E}_{y \sim P_{\text{gen}}(y)} [\log(1 - D(G(y)))]. \quad (1)$$

此处 G 为生成器, D 为判别器, P_{train} , P_{gen} 分别代表训练数据和要生成数据的分布, 而 x , y 分别代表水下原始图像和水下参考图像。判别器的数据输入来自参考图像和由生成器生成的图像。这两类数据分别与原始图像串联作为输入, 将判别目标从二元分类问题转变为目标图像类相似度估计问题。它可以防止生成器生成异常图像以获得模糊的判别结果。在训练过程中, 生成器和判别器均保持独立训练, 以防止它们之间的协作优化。同时使二者交替训练以确保梯度平稳。当判别器和生成器的损失稳定地波动在某一常量时, 则二者达到动态平衡, 进而通过生成器可以获得清晰的水下图像。此外, 为了提高训练的稳定性, 放宽了判别准则, 即将判别为假的目标值设为 0~0.1 之间的随机数, 而把判别真的目标值设为 0.9~1.05, 以减少异常数据引起的不稳定性。

考虑到输出图像中的人为干扰, 使用 L1 损失解决此问题, 这有助于在训练的早期阶段迅速稳定网络。生成图像 $G(x)$ 和参考图像 y 之间的 L1 损失函数如公式(2)所示:

$$\mathcal{L}_{L_1}(G) = \mathbb{E}[\|y - G(x)\|_1]. \quad (2)$$

MSE 损失表示生成的图像与原始图像之间的均方误差。生成的图像 $G(x)$ 和参考图像 y 的 MSE 损失如公式(3)所示:

$$\mathcal{L}_{\text{MSE}}(G) = \frac{1}{\text{CWH}} [\|y - G(x)\|_2^2]. \quad (3)$$

其中 C、W、H 为常数, 分别代表图像的通道数, 宽度和高度。

结构相似性(SSIM) [23]可以根据亮度、对比度和结构这三个属性对图像进行比较。其描述增强后的图像和参考图像在上述三个指标中的相似度, 数值范围介于 0 到 1。并且数值越大, 与参考图像的相似度越高。该指标更接近于人类的直觉, 可以得出与主观评价一致的结论。在损失函数中加入 SSIM 损失来增强结果, 如公式(4)所示:

$$\mathcal{L}_{\text{SSIM}}(x, y) = 1 - \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}. \quad (4)$$

在公式(4)中, μ_x 、 μ_y 表示平均值, 而 μ_x^2 、 μ_y^2 分别是 x 和 y 的方差; σ_{xy} 表示 x 和 y 之间的相关性。两个常数 $c_1 = (255 \times 0.01)^2$, $c_2 = (255 \times 0.03)^2$, 用于确保除数为零时的数值稳定性。

另一种构建内容损失的方法是使用特征损失, 其目的是比较两幅图像之间在特征空间的差异。特征损失函数用于水下图像的特征重建, 以减少图像的模糊性, 获得轮廓结构清晰的结果。我们定义特征损失 \mathcal{L}_F 如公式(5)所示:

$$\mathcal{L}_F = \frac{1}{\text{CWH}} \|F(y) - F(G(x))\|_2^2, \quad (5)$$

其中 F 表示非线性 CNN 变换, 特征图大小为 $W \times H \times C$ 。该算法使用预先训练好的 VGG19 模型作为特征提取器, 并将卷积层的特征图在第 3 个池化层(conv3_3)之前进行损失计算。

综上, 通过线性连接上述公式(1)~(5), 可以得到总损失函数如下:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_F + \lambda_1 \mathcal{L}_{L_1} + \lambda_2 \mathcal{L}_{\text{MSE}} + \lambda_3 \mathcal{L}_{\text{SSIM}} + \lambda_4 \mathcal{L}_{\text{GAN}}. \quad (6)$$

4. 实验结果与分析

本文使用公开的数据集进行实验, 并在相同条件下与其他算法(包括基于深度学习的算法和传统算法)

进行比较。然后, 我们对测试图像进行了主观和客观的评估, 最后用两种典型的应用算法验证我们方法的实用性。事实证明, 该方法具有良好的增强效果。

4.1. 数据集与训练细节

本文中用于网络的水下图像数据集来自 EUVP [10]数据集的 imagenet 子集, 该子集包含数据中广泛的自然变化(例如场景、水类型、照明条件等)。在本文中, 3600 张图像被用作训练集, 100 张图像被用作测试集。网络的输入图像大小被限制长和宽为 256 的 RGB 图像。优化器为 Adam, alpha 值设置为 0.9, beta 值设置为 0.5, 学习率为 0.0001, 经过一半的训练后, 它被缩小到原始大小的一半。批量大小(batch size)为 6, 周期(epoch)为 60。参考[8]和[13]中组合损失的初始比值, 将损失函数中的权重 λ_1 、 λ_2 、 λ_3 、 λ_4 分别设置为 0.14、0.06、0.6 和 0.2。对于所有其他基于深度学习的方法, 我们使用其默认参数值, 并且在相同的训练和测试集上进行训练和测试。

4.2. 主观评价

所提出的算法与基于对比度约束的自适应直方图均衡(CLAHE) [24]、水下图像融合增强(FE) [7]、基于水下光衰减先验(ULAP) [5]、基于弱监督色彩转换(WRCT) [12]和基于生成式对抗网络的水下图像增强(UGAN) [13]进行比较。

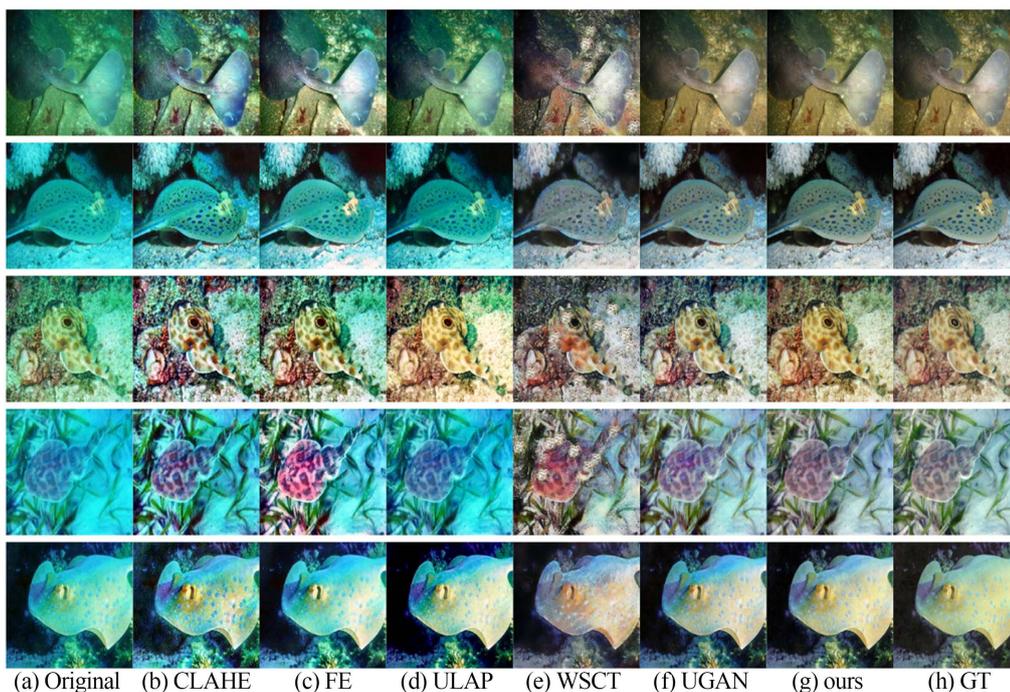


Figure 5. Comparison of experimental results of different underwater image enhancement algorithms
图 5. 多尺度密集块结构

对比实验结果显示在图 5 中。CLAHE 算法的恢复结果具有较高的对比度, 但仍出现偏色的色调。由于颜色校正算法不正确, FE 算法有明显的红色偏色。例如, 第四幅图像中的深棕色被恢复为亮红色。ULAP 算法对光线充足的图像比较好, 但对光线不足的场景, 图像的恢复效果很差。前两种算法在恢复蓝绿色调时都有困难。WSCT 算法对比度高, 在一些水下图像中引入了不相干的噪声, 导致水下图像的原始内容被改变, 主要由于该方法没有使用成对的训练图像。例如, 在第一幅图像中, 使用 WSCT 算法将背景

中的白色岩石还原为相应的白噪声, 而最后一幅图像中的圆形斑点被还原为其他形状。UGAN 在提高对比度和清晰度方面表现良好。相比之下, 本文算法在去除噪声(如第三张)和颜色信息校正(如第四张和第五张)方面具有优势, 在视觉上更具有吸引力。

从图 5 中第二张背景为沙子和珊瑚的图像中可以看出, 对于远处区域的高频信息, 本文算法的恢复结果细节完整, UGAN 在沙石区域细节缺失, WSCT 则有明显的模糊现象。这是由于使用特征损失可以得到很好的细节, 使恢复结果更加清晰。没有使用生成网络的传统方法无法从高频细节损失的不可逆操作中再生出高频细节。

4.3. 客观评价

将本文所提算法和其他算法使用不同指标进行比较。我们使用均方误差(MSE)、峰值信噪比(PSNR)、结构相似度(SSIM)、水下图像质量测量指标(Underwater Image Quality Measure, UIQM) [25]和自然图像质量评价指标(Natural Image Quality Evaluator, NIQE) [26]对各个算法的测试结果进行了定量比较。MSE 指标描述了由增强算法获得的图像与参考图像在像素层面上的差异。该值越小, 两幅图像的像素值就越接近。PSNR 指标描述了生成的图像中保留了多少原始信息。数值越大, 图像中保留的原始信息就越多。SSIM 指标描述了两幅图像在亮度和对比度等主观指标方面的相似度。该值越大, 说明图像与参考图像在高级特征方面的重叠程度越高。UIQM 代表水下图像质量指标, 它只需要一张图像, 是一个非参考评价指标, 更适合于开放的水下任务。UIQM 的评价分为三部分, 即图像颜色指标(UICM)、清晰度指标(UISM)和对比度指标(UICoM)。UIQM 的计算方法如公式(7)所示:

$$UIQM = c_1 \times UICM + c_2 \times UISM + c_3 \times UICoM, \quad (7)$$

其中 c_1 、 c_2 和 c_3 是与应用相关的参数。NIQE 是另一个无参考评价指标, 它描述的是目标图像特征模型距离预设的标准自然图像特征模型之间的差距, 与水下图像增强目标相契合。NIQE 的值与图像的自然状态呈现负相关。该值越小, 图像在主观上被认为越接近自然状态。NIQE 如公式(8):

$$D(v_1, v_2, \Sigma_1, \Sigma_2) = \sqrt{\left((v_1 - v_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (v_1 - v_2) \right)}, \quad (8)$$

v_1 、 v_2 表示自然图像和待评价图像在 MVG 特征模型之间的均值向量, Σ_1 、 Σ_2 则表示上述两张图像在 MVG 特征模型间的协方差矩阵。

表 1 为各种算法在测试集上的比较结果与标准差, 加粗字体为最优值。从表 1 中可以看到, CLAHE、FE 和 ULAP 方法在 MSE、PSNR 这种硬性指标下表现都不好。WSCT 方法具有较好的 UIQM 值, 主要由于 UIQM 指标中 UICoM 值的比重较高, 其更加希望生成具有较高对比度的图像。UGAN 方法和本文方法在各项指标都要优于其他方法。本文算法得到的结果无论在图像质量和稳定性方面都要普遍优于 UGAN。

Table 1. Quantitative comparison of different underwater image enhancement algorithms

表 1. 不同水下图像增强算法在测试集上的定量比较

Method	MSE	PSNR	SSIM	UIQM	NIQE
CLAHE [24]	1994.1 ± 884.4	15.4957 ± 1.7254	0.6392 ± 0.0540	2.8549 ± 0.4038	4.9615
FE [7]	1706.0 ± 965.9	16.3661 ± 2.1262	0.6716 ± 0.1053	2.8794 ± 0.3221	5.3086
ULAP [5]	3836.0 ± 2582.6	13.3359 ± 3.1285	0.5584 ± 0.1035	1.9750 ± 0.5613	5.0077
WSCT [12]	842.3 ± 583.7	19.8133 ± 2.9057	0.5657 ± 0.1160	3.3365 ± 0.1328	4.7094
UGAN [13]	563.2 ± 543.5	21.8858 ± 3.1040	0.7555 ± 0.0570	3.2359 ± 0.1572	4.6669
Ours	341.5 ± 281.5	23.8260 ± 2.8326	0.7983 ± 0.0501	3.2365 ± 0.2150	4.5902

5. 总结

水下图像受特殊成像环境影响存在颜色失真、对比度低等问题, 本文提出一种水下图像增强算法, 该算法在 GAN 网络中加入多尺度密集块和注意力机制块。在生成器网络中使用残差多尺度密集块, 提高了网络对特征的提取能力, 并且加入两组空间-通道注意力机制块, 增强了网络对复杂特征的筛选能力, 提升了网络的收敛速度。同时将多种内容损失和结构损失相结合, 在不同尺度空间对网络进行约束, 提升了 GAN 网络的稳定性和对不同图像的泛用性。最后进行了多种对比试验, 证明本文算法可以对失真水下图像进行较好的增强, 提升图像的整体视觉效果。

参考文献

- [1] Mobley, C.D. (1994) *Light and Water: Radiative Transfer in Natural Waters*. Academic Press, Cambridge.
- [2] Xiong, W., et al. (2019) Foreground-Aware Image Inpainting. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 16-20 June 2019, 5833-5841. <https://doi.org/10.1109/CVPR.2019.00599>
- [3] Morel, A. (1978) Marine Optics. *Earth Science Reviews*, **14**, 170-171. [https://doi.org/10.1016/0012-8252\(78\)90007-7](https://doi.org/10.1016/0012-8252(78)90007-7)
- [4] Jae, J.S. (1990) Computer Modeling and the Design of Optimal Underwater Imaging Systems. *IEEE Journal of Oceanic Engineering*, **15**, 101-111. <https://doi.org/10.1109/48.50695>
- [5] Song, W., et al. (2018) A Rapid Scene Depth Estimation Model Based on Underwater Light Attenuation Prior for Underwater Image Restoration. *Advances in Multimedia Information Processing*, Hefei, 21-22 September 2018, 678-688. https://doi.org/10.1007/978-3-030-00776-8_62
- [6] Huang, D., Yan, W., Wei, S., et al. (2018) Shallow-Water Image Enhancement Using Relative Global Histogram Stretching Based on Adaptive Parameter Acquisition. In: *International Conference on Multimedia Modeling*, Springer, Cham, 453-465. https://doi.org/10.1007/978-3-319-73603-7_37
- [7] Ancuti, C., Ancuti, C.O., Haber, T. and Bekaert, P. (2012) Enhancing Underwater Images and Videos by Fusion. 2012 *IEEE Conference on Computer Vision and Pattern Recognition*, Providence, 16-21 June 2012, 81-88. <https://doi.org/10.1109/CVPR.2012.6247661>
- [8] Li, C., et al. (2020) An Underwater Image Enhancement Benchmark Dataset and Beyond. *IEEE Transactions on Image Processing*, **29**, 4376-4389. <https://doi.org/10.1109/TIP.2019.2955241>
- [9] Hou, M., Liu, R., Fan, X. and Luo, Z. (2018) Joint Residual Learning for Underwater Image Enhancement. 2018 *25th IEEE International Conference on Image Processing (ICIP)*, Athens, 7-10 October 2018, 4043-4047. <https://doi.org/10.1109/ICIP.2018.8451209>
- [10] Islam, M.J., Luo, P. and Sattar, J. (2020) Simultaneous Enhancement and Super-Resolution of Underwater Imagery for Improved Visual Perception.
- [11] Uplavikar, P., Wu, Z. and Wang, Z. (2019) All-in-One Underwater Image Enhancement Using Domain-Adversarial Learning.
- [12] Li, C., Guo, J. and Guo, C. (2017) Emerging from Water: Underwater Image Color Correction Based on Weakly Supervised Color Transfer. *IEEE Signal Processing Letters*, **25**, 323-327.
- [13] Fabbri, C., Islam, M.J. and Sattar, J. (2018) Enhancing Underwater Imagery Using Generative Adversarial Networks. 2018 *IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, 21-25 May 2018, 7159-7165. <https://doi.org/10.1109/ICRA.2018.8460552>
- [14] Trucco, E. and Olmos-Antillon, A.T. (2006) Self-Tuning Underwater Image Restoration. *IEEE Journal of Oceanic Engineering*, **31**, 511-519. <https://doi.org/10.1109/JOE.2004.836395>
- [15] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. https://doi.org/10.1007/978-3-319-24574-4_28
- [16] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 640-651. <https://doi.org/10.1109/TPAMI.2016.2572683>
- [17] Szegedy, C., et al. (2015) Going Deeper with Convolutions. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
- [18] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>

- [19] Guo, Y., Li, H. and Zhuang, P. (2019) Underwater Image Enhancement Using a Multiscale Dense Generative Adversarial Network. *IEEE Journal of Oceanic Engineering*, **45**, 862-870.
- [20] Liu, F., Shen, C., Lin, G., *et al.* (2015) Learning Depth from Single Monocular Images Using Deep Convolutional Neural Fields. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **38**, 2024-2039. <https://doi.org/10.1109/TPAMI.2015.2505283>
- [21] Miyato, T., Kataoka, T., Koyama, M., *et al.* (2018) Spectral Normalization for Generative Adversarial Networks. *International Conference on Learning Representations*, Vancouver, 30 April-3 May 2018.
- [22] Li, H., Li, J. and Wang, W. (2019) A Fusion Adversarial Underwater Image Enhancement Network with a Public Test Dataset.
- [23] Wang, Z., Bovik, A.C., Sheikh, H.R., *et al.* (2004) Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, **13**, 600-612. <https://doi.org/10.1109/TIP.2003.819861>
- [24] Zuiderveld, K. (1994) Contrast Limited Adaptive Histogram Equalization. In: *Graphics Gems*, Morgan Kaufmann Publishers, Burlington, 474-485. <https://doi.org/10.1016/B978-0-12-336156-1.50061-6>
- [25] Panetta, K., Gao, C. and Agaian, S. (2016) Human-Visual-System-Inspired Underwater Image Quality Measures. *IEEE Journal of Oceanic Engineering*, **41**, 541-551. <https://doi.org/10.1109/JOE.2015.2469915>
- [26] Mittal, A., *et al.* (2013) Making a “Completely Blind” Image Quality Analyzer. *IEEE Signal Processing Letters*, **20**, 209-212. <https://doi.org/10.1109/LSP.2012.2227726>