

# 基于金字塔注意力机制和生成对抗网络的水下图像增强

王越, 王德兴\*, 袁红春\*\*, 吴若有, 官鹏

上海海洋大学信息学院, 上海 201306

**摘要** 为提升水下图像的增强效果,提出了一种基于金字塔注意力机制和生成对抗网络(GAN)的水下图像增强算法,该算法将生成对抗网络作为基本架构,生成网络采用编码解码结构并引入特征金字塔注意力模块,多尺度金字塔特征与注意力机制的结合有利于捕获更丰富的高级特征以提升模型性能;判别网络采用类似马尔可夫判别器的结构。此外,通过构建包含全局相似性、内容感知和色彩感知的多项损失函数,使增强后的图像与参考图像的结构、内容和色彩保持一致。实验结果表明,所提算法增强的水下图像在清晰度、颜色校正和对比度上都有所提升。其中,结构相似性、水下图像质量度量 and 信息熵的平均值分别为 0.7418、2.9457 和 4.6925。在主观感知和客观评价指标上,所提算法的实验结果均优于对比算法。

**关键词** 图像处理; 水下图像增强; 注意力机制; 生成对抗网络; 编码解码结构

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.1610006

## Underwater Image Enhancement Based on Pyramid Attention Mechanism and Generative Adversarial Network

Wang Yue, Wang Dexing\*, Yuan Hongchun\*\*, Wu Ruoyou, Gong Peng

School of Information, Shanghai Ocean University, Shanghai 201306, China

**Abstract** This study proposes an underwater image enhancement algorithm based on the pyramid attention mechanism and generative adversarial network (GAN) to improve the enhancement effect of underwater images. It uses the generative adversarial network as the basic architecture, and the generative network adopts the encoding and decoding structures and introduces the feature pyramid attention module. The combination of multi-scale pyramid features and attention mechanism can capture richer advanced features to improve model performance, and the structure of the discriminant network is similar to the Markov discriminator. In addition, a multi-loss function including global similarity, content perception, and color perception is constructed to keep the structure, content, and color of the enhanced image consistent with those of the reference image. The experimental results show that the sharpness, color correction, and contrast of underwater images enhanced by the proposed algorithm are improved. The average values of the structural similarity, underwater image quality measurement, and information entropy are 0.7418, 2.9457, and 4.6925, respectively. For subjective perception and objective evaluation indicators, the experimental results of the proposed algorithm are better than that of the comparison algorithm.

**Key words** image processing; underwater image enhancement; attention mechanism; generative adversarial network; encoding and decoding structure

**OCIS codes** 100.2980; 100.4996; 100.3010

收稿日期: 2020-10-16; 修回日期: 2020-11-25; 录用日期: 2020-12-08

基金项目: 国家自然科学基金(41776142)

通信作者: \*dxwang@shou.edu.cn; \*\*hcyuan@shou.edu.cn

# 1 引言

随着陆地上可用资源的减少,海洋资源的开发和利用成为人类亟需面对的问题。近年来,水下机器人被广泛应用于海洋资源探索等方面。由于水下环境较为复杂,水下机器人拍摄到的图像通常出现严重的退化,主要呈现为模糊、对比度低、颜色偏差等,这使得水下机器人无法获取高质量的清晰水下图像,严重影响了水下图像在海底测绘、海洋物种迁徙监测<sup>[1]</sup>、水下目标检测和识别<sup>[2]</sup>等领域的应用,因此对水下图像进行增强处理具有重要意义。

目前,水下图像增强方法分为三种:基于非物理模型的方法,基于物理模型的方法和基于数据驱动的方法<sup>[3]</sup>。基于非物理模型的方法通过调整图像像素取值实现图像的增强<sup>[4-7]</sup>,从而提升视觉感知。例如,Iqbal 等<sup>[5]</sup>拉伸了 RGB 颜色空间和 HSV 颜色空间中的动态像素范围,以提高水下图像的饱和度和对比度,该方法的计算代价高,而且会对图像的局部区域造成过度增强或增强不足的问题。基于物理模型的方法依据物理成像模型,将水下图像增强看作是水下成像过程的逆问题,通过逆求解水下成像模型复原高质量的水下图像<sup>[8-11]</sup>。该方法需要利用水下图像的物理先验估计增强模型的参数,对模型参数的估计存在一定的偏差,因此对图像的增强效果不佳。文献[9]将波长相关补偿算法与图像去雾算法相结合以复原水下图像,能有效地消除水下图像的蓝绿色调和人造光源的影响。而该方法对色偏严重的水下图像进行处理时,不能很好地消除图像的颜色偏差。基于数据驱动的方法主要分为两大类:基于卷积神经网络(CNN)的方法<sup>[12-13]</sup>与基于生成对抗网络(GAN)的方法<sup>[14-15]</sup>。特别是基于卷积神经网络的残差模型<sup>[16]</sup>和基于生成对抗网络的模型,在去除阴影、增强对比度等方面表现良好。然而,一些增强方法的性能在颜色校正方面还有待提升。Islam 等<sup>[17]</sup>设计了一个简单的生成对抗网络模型以实时增强水下图像,利用该方法增强后的水下图像会引入意想不到的颜色,图像颜色失真,且整体偏暗。近年来,金字塔注意力机制被成功地应用于语义分割领域<sup>[18]</sup>,通过金字塔结构和注意力机制的结合,可以捕获更丰富的高级特征,更好地学习图像细节,从而有效提升模型的性能。这一应用在水下图像增强领域中还未实现。

针对现有方法所存在的缺陷,本文提出了一种基于金字塔注意力机制和生成对抗网络的水下图像

增强算法。该算法以 GAN 为基本框架,利用生成网络学习退化水下图像与清晰水下图像间的映射关系。生成网络中引入了特征金字塔注意力模块,通过融合不同尺度的上下文信息,为高层次的特征图提供更好的像素级注意力。而判别网络用来学习、区分生成图像与其对应的参考图像。此外,通过在损失函数中加入全局相似性、内容感知和色彩感知损失函数,可进一步提升增强模型的性能。实验结果表明,本文算法能有效地校正颜色偏差,增强对比度,提高图像的视觉感知。

## 2 基本原理

### 2.1 卷积层

卷积神经网络和生成对抗网络中的卷积层可提取输入的不同特征,并对输入的特征图进行卷积操作,每层卷积层的输出特征图的计算公式为

$$f^{i+1} = \text{Activation}(f^i * k^{i+1} + b^{i+1}), \quad (1)$$

式中: $k^{i+1}$  代表卷积核; $b^{i+1}$  代表偏差;Activation( $\cdot$ )代表激活函数; $f^i, f^{i+1}$  分别对应着第  $i$  层和第  $i+1$  层卷积层的输出特征图。

每层卷积层的输出特征图的尺寸为

$$O = \left\lfloor \frac{n + 2 * p + k}{s} \right\rfloor + 1, \quad (2)$$

式中: $O$  代表输出特征图的尺寸; $n$  代表输入特征图的尺寸; $k$  代表卷积核的尺寸; $p$  代表零填充数; $s$  代表步长。当  $s$  的值等于 2 时,该卷积操作相当于步长为 2 的池化操作,得到的输出特征图尺寸为输入特征图尺寸的一半。

### 2.2 特征金字塔注意力模块

特征金字塔注意力(FPA)<sup>[18]</sup>模块实现了利用一个类似 U 形的结构提取三种不同尺度的金字塔特征的功能,且该模块能够融合来自三种不同尺度下相邻的金字塔特征。金字塔结构中分别使用了  $3 \times 3, 5 \times 5, 7 \times 7$  的卷积核,以便于更好地从不同尺度下的金字塔特征中提取上下文信息。由于网络输出的高层次特征图的尺寸通常较小,使用较大尺寸的卷积核不会引入过多的计算代价。金字塔结构逐步整合不同尺度的特征信息,以便更准确地融合三个相邻尺度的上下文特征。通过  $1 \times 1$  的卷积操作,将一系列卷积层提取的原始特征与经过金字塔结构处理的特征进行逐像素相乘。此外,引入了全局池化分支与前述相乘后的特征进行联结操作,这进一步提高了 FPA 模块的性能。FPA 模块将空间金字塔结构与注意力机制相结合以提取准确而密集的特

征,其生成的每层特征图都是信息加强的图像,且为高层次的特征图提供了更好的像素级注意力,从而有效提升了模型的性能。此外,该模块融合的上下文信息和卷积层提取的原始特征是逐像素相乘的,因此该运算的计算量不会太大。最终的模块结构如图 1 所示。

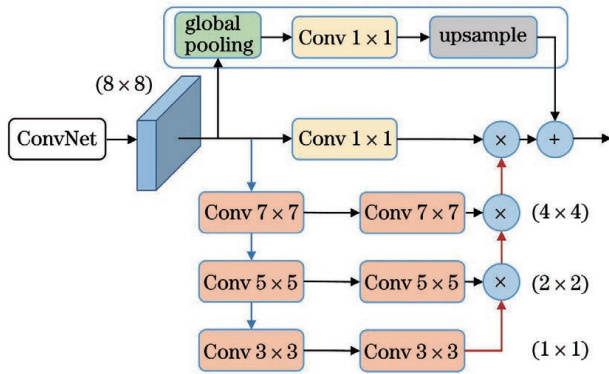


图 1 特征金字塔注意力模块  
Fig. 1 Feature pyramid attention module

### 3 FPAGAN 增强模型

本文提出了一种基于金字塔注意力机制和生成对抗网络的水下图像增强算法。该算法将配对的水

下图像作为训练数据,并构建了多项损失函数来提升网络模型输出的图像效果,通过对网络模型的训练,最终得到对比度提高、颜色鲜艳、清晰的水下图像。

#### 3.1 模型结构

本文提出了一个网络模型 FPAGAN (generative adversarial network with feature pyramid attention), 利用编码器与解码器提取图像特征,并与特征金字塔注意模块相结合,对退化的水下图像进行增强。FPAGAN 的结构分为生成网络和判别网络两个部分。第一部分是生成网络,目标是学习退化水下图像与清晰水下图像间的映射关系,以生成高质量的清晰水下图像。第二部分是判别网络,目标是使网络学会区分生成的虚假图像和真实的参考图像,以促使生成的图像更接近于参考图像。本文算法的执行流程如下:先将退化的水下图像作为生成网络的输入,经过模型的推理得到生成图像;将生成图像和参考图像一起送到判别网络中,计算出两张图像间的误差损失;再将计算出的误差损失送回生成网络的各个层,以对网络的各层参数进行不断更新,从而获得高质量的、视觉感知良好的水下图像。FPAGAN 的网络结构以及数据处理流程如图 2 所示。

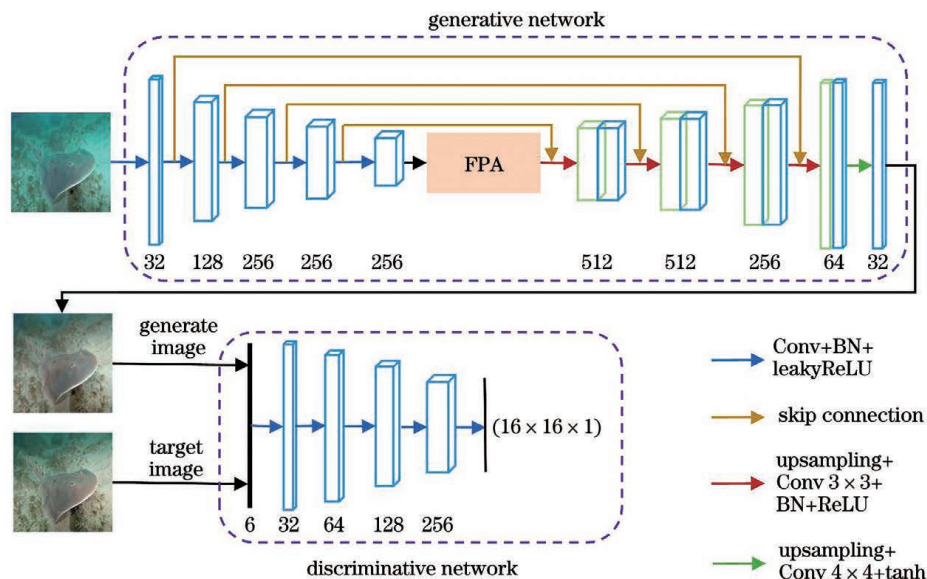


图 2 FPAGAN 结构  
Fig. 2 Structure of FPAGAN

#### 3.2 生成网络的构建

受 Unet<sup>[19]</sup> 网络的启发,生成网络包含编码模块、FPA 模块和解码模块三个主要部分,其网络结构如图 3 所示。图 3 中,第一个卷积层的 5×5、32、2、same 分别代表卷积核大小、卷积核数量、步长以

及补零填充策略。BN(batch normalization)表示批量归一化。ReLU、tanh、leakyReLU 分别代表不同的非线性激活函数。

在生成网络中使用步长为 2 的步幅卷积代替传统的池化层,以避免传统池化操作的局限性。此



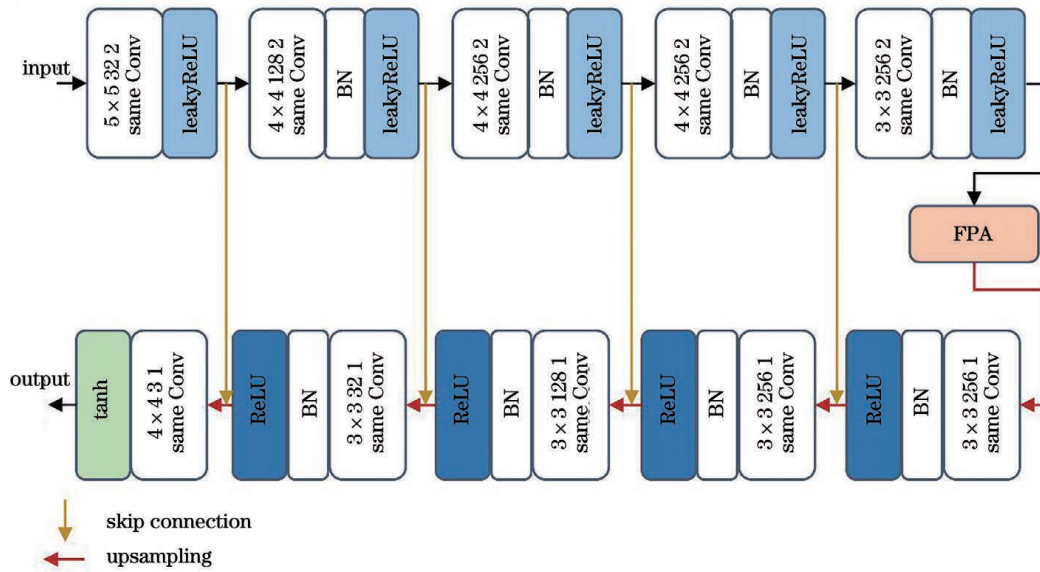


图 3 生成网络结构  
Fig. 3 Generative network structure

外,生成网络中添加了跳跃连接操作,以避免图像的细节损失。如图 3 所示,首先利用生成网络中的编码模块对输入图像进行下采样操作,输出尺寸为  $8 \times 8 \times 256$  的特征图;然后将输出的特征图经过 FPA 模块的处理,得到包含不同尺度上下文信息的特征图;最后,利用生成网络中的解码模块对特征图进行上采样,得到尺寸为  $256 \times 256 \times 3$  的生成图像。

### 3.3 判别网络的构建

判别网络的结构如图 4 所示。判别网络采用了

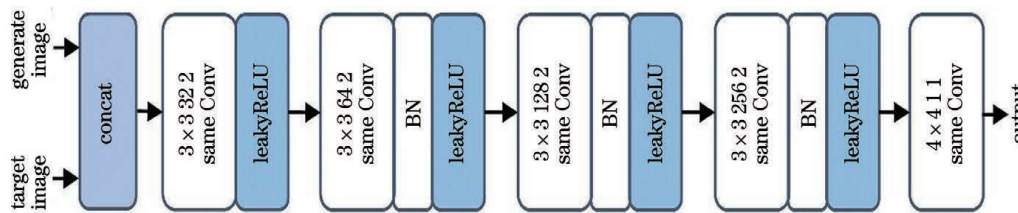


图 4 判别网络结构  
Fig. 4 Discriminative network structure

在判别网络中,首先使用 concat 操作将生成图像与参考图像按通道维度叠加,得到尺寸为  $256 \times 256 \times 6$  的图像;然后将叠加后的图像通过一系列卷积、批量归一化、非线性映射操作,最终得到  $16 \times 16 \times 1$  大小的输出,该输出表示判别网络对生成图像和参考图像之间差异的平均响应。

### 3.4 损失函数的构建

本文通过对生成网络和判别网络的对抗训练来优化 FPAGAN 模型,网络的优化过程对应着损失函数的优化。对抗网络的损失函数为

5 个卷积层,类似于马尔可夫判别器(PatchGAN)<sup>[20]</sup>体系结构,PatchGAN 仅基于补丁级别信息进行区分。判别网络以生成网络生成的图像和参考图像作为输入,输出为  $16 \times 16 \times 1$  大小的补丁块。如果生成图像和参考图像的某个区域差距较小,则输出补丁块中对应的位置像素点的值接近于 1,反之接近于 0。它的优点是输出补丁块的每一个像素点对应输入图像的一个局部感受野,进而可判断输入图像局部感受野之间的差别,而不是整幅图像之间的差别,最终减少了人为的干扰因素。

$$L_{adv} = E(y) \{ \log [D(y)] \} + E(x) \{ \log \{ 1 - D[G(x)] \} \}, \quad (3)$$

式中: $D$  表示判别网络; $G$  表示生成网络; $x$  表示退化的水下图像; $y$  表示与退化图像对应的参考图像; $E$  表示数学期望。

模型训练过程中,对判别网络更新一次再对生成网络更新一次,以防止模型梯度消失。如此循环反复进行参数更新,当判别网络不能区分生成图像和参考图像时,模型达到动态平衡,此时便可获得清晰的水下图像。为了减少输出图像中的人为干扰因

素,在本文中考虑了三个方面,即全局相似性、图像内容以及色彩感知损失。

全局相似性损失:现有的方法表明,在目标函数中添加  $L_1(L_2)$  损失可以使生成网络  $G$  学习图像之间的全局相似性<sup>[21]</sup>。由于  $L_1$  损失不易引入模糊,在目标函数中加入了  $L_1$  损失项:

$$L_1 = E_{x,y} [\|y - G(x)\|_1], \quad (4)$$

式中: $x$  表示退化的水下图像; $y$  表示与退化图像对应的参考图像。

内容感知损失:图像内容感知损失项可以驱动  $G$  生成与参考图像内容相似的增强图像。受文献<sup>[22-23]</sup>的启发,将图像内容函数  $\varphi(\cdot)$  定义为由预先训练的 VGG-19 网络的 conv5\_2 层提取的高级特征。内容感知损失的定义为

$$L_{\text{con}} = E_{x,y} \{ \|\varphi(y) - \varphi[G(x)]\|_2 \}, \quad (5)$$

式中: $\varphi(\cdot)$  表示图像内容函数。

色彩感知损失:该损失项可促进  $G$  生成与参考图像色彩相似的增强图像,其定义如下:

$$L_{\text{col\_per}} = \sqrt{\frac{(512 + r_{\text{mean}})r^2}{256} + 4g^2 + \frac{(767 - r_{\text{mean}})b^2}{256}}, \quad (6)$$

式中: $r$ 、 $g$ 、 $b$  分别表示参考图像与生成图像的红、绿、蓝颜色通道的差值, $r_{\text{mean}}$  表示参考图像与生成图像红色通道的平均值,可分别表示为

$$r_{\text{mean}} = \frac{y_{c_r} + G(x)_{c_r}}{2}, \quad (7)$$

$$r = y_{c_r} - G(x)_{c_r}, \quad (8)$$

$$g = y_{c_g} - G(x)_{c_g}, \quad (9)$$

$$b = y_{c_b} - G(x)_{c_b}, \quad (10)$$

式中:下标  $c_r$ 、 $c_g$ 、 $c_b$  分别表示图像的红、绿、蓝三颜色通道。

FPAGAN 使用多项损失函数进行网络模型的学习,这有利于网络参数的加速收敛,同时可以提高模型的鲁棒性。目标函数通过将对抗损失  $L_{\text{adv}}$ 、 $L_1$  损失函数、内容感知损失  $L_{\text{con}}$  和色彩感知损失  $L_{\text{col\_per}}$  进行线性叠加,得到

$$L = L_{\text{adv}} + \lambda_1 L_1 + \lambda_2 L_{\text{con}} + \lambda_3 L_{\text{col\_per}}, \quad (11)$$

式中: $\lambda_1$ 、 $\lambda_2$ 、 $\lambda_3$  代表比例因子,根据经验分别设置为 0.6、0.3、0.1。

### 3.5 实验设置

本实验在 Windows 操作系统下,基于 Tensorflow 和 Keras 深度学习开源框架实现。计算机的配置如下: NVIDIA GeForce RTX2070 Max-Q (8 GB)

GPU, Inter Core i7-10750H CPU, 主频 2.60 GHz, 内存 16 GB。

本文实验中将所有训练样本的尺寸缩小至  $256 \times 256 \times 3$ , 并将其归一化到  $[-1, 1]$  区间; batch\_size 为 8; 网络模型中共采用了 5 种不同尺寸的卷积核, 大小分别为  $7 \times 7$ 、 $5 \times 5$ 、 $4 \times 4$ 、 $3 \times 3$ 、 $1 \times 1$ ; 此外, 采用了 3 种不同的非线性激活函数以提高模型的泛化能力, 分别为 ReLU、tanh、leakyReLU; 为了降低深度神经网络的训练难度, 在部分卷积操作后加入了批量归一化 BN, 它能加速模型的收敛、减小振荡、使训练过程更加稳定, 同时可以防止训练后的模型出现过拟合。模型训练过程中采用 Adam 优化器 ( $\beta_1$  为 0.5,  $\beta_2$  为 0.999); 设置生成网络和判别网络的初始学习率为 0.0001, 训练轮数 epoch 为 50。

## 4 分析与讨论

### 4.1 实验数据制作

现阶段, 获取大量真实世界水下图像的参考图像比较困难, 这对分析模型学习退化的水下图像与其参考图像间的映射关系形成了一定的挑战。为了解决缺乏成对的训练数据问题, Islam 等<sup>[17]</sup> 提出了大规模真实世界水下图像数据集 EUVP, 利用现有的数据集 EUVP 构造了一个用于训练 FPAGAN 模型的数据集 EUID。该数据集共有 8739 张图像, 它被划分为训练集、验证集和测试集, 其中包含作为训练集的 3700 对配对的水下图像, 作为验证集的 1270 张水下图像, 作为测试集的 69 张水下图像。其中, 测试集又被划分为测试集 A 和测试集 B, 测试集 A 包含 46 张有参考的水下图像, 测试集 B 包含 23 张无参考的水下图像。验证集一般用在训练的过程中, 每个 epoch 结束后可以在验证集上验证增强效果, 目的是及时发现模型或者参数的问题和验证模型的泛化能力, 从而可以及时做出应对措施。而测试集一般用在训练结束之后, 目的是评估最终训练好的模型性能。

### 4.2 实验分析

为验证本文算法的有效性, 通过大量实验, 将其与现有的水下图像增强方法进行对比, 包括基于非物理模型的方法 (GC)、基于物理模型的方法 (UDCP, LDGP)、基于数据驱动的方法 (UWCNN, FUnIE-GAN)。本文中首先对引入的 FPA 模块的有效性进行分析, 其次分别在测试集 A 和测试集 B 上对以上算法的实验结果进行定性和定量对比。

#### 4.2.1 FPA 模块的对比实验

为了证明在生成网络中加入的 FPA 模块的有效性,对本文模型和未采用 FPA 模块的 GAN 模型

进行对比实验。其中,两种模型在训练集及其超参数的取值等方面均保持一致。在测试集上的主观实验结果如图 5 所示。

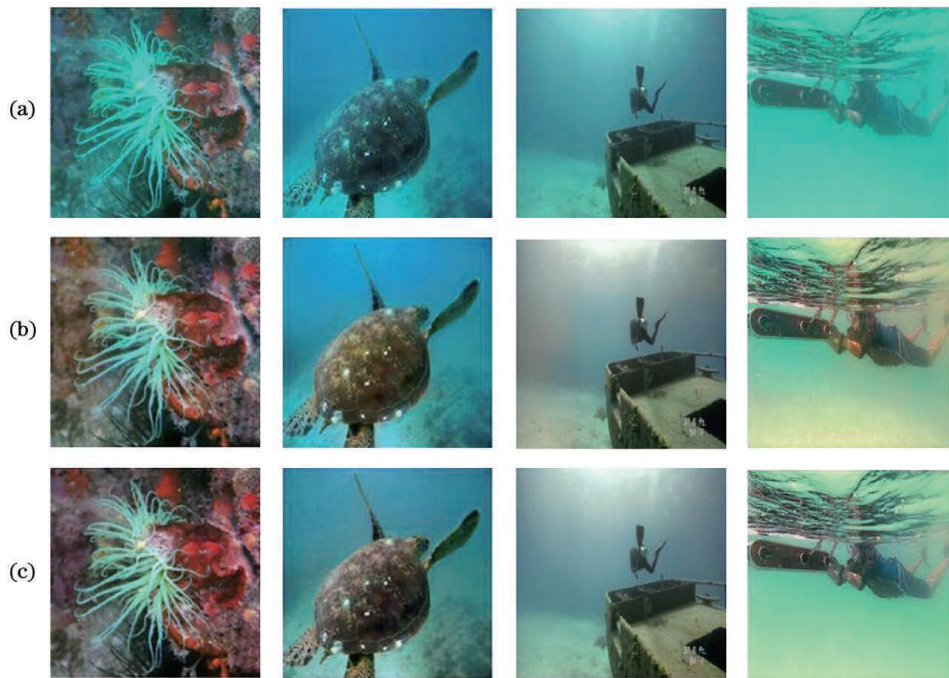


图 5 未采用 FPA 模块和采用 FPA 模块的对比实验结果。(a)水下图像;(b)GAN;(c)FPAGAN

Fig. 5 Results of comparative experiments without FPA module and with FPA module. (a) Underwater images; (b) GAN; (c) FPAGAN

由图 5 可知,未采用 FPA 模块的 GAN 模型能够有效地校正呈现蓝绿色调的水下图像,但同时引入了淡黄色,图像整体偏暗,增强效果不佳。本文算法将空间金字塔结构与注意力机制相结合,提取了图像准确而密集的特征和上下文信息,有效地提升了图像的视觉感知,校正了颜色偏差,增强了对比度。

为了进一步验证 FPA 模块的有效性,对两种模型在测试集 A 和测试集 B 上的性能进行定量分析。首先,考虑了两种全参考评价指标<sup>[23-25]</sup>,即峰值信噪比(PSNR)和结构相似性(SSIM)。其中,PSNR 值和 SSIM 值越大,待评价图像的亮度、对比度、整体结构以及包含的信息越接近参考图像。其次,考虑了三种非参考图像质量评价指标——水下图像质量度量(UIQM)<sup>[26]</sup>、信息熵(IE)、自然图像质量评估(NIQE)<sup>[27]</sup>。定量分析结果如表 1 和表 2 所示。

通过表 1 中的实验结果能够发现,加入 FPA 模块后,增强后的水下图像与参考图像之间的失真程度更小,图像的亮度、对比度、结构信息等方面更接近于参考图像。

表 1 GAN 和 FPAGAN 在测试集 A 上的实验结果

Table 1 Experimental results obtained by GAN and FPAGAN on test set A

Method	PSNR	SSIM
GAN	21.9328	0.7330
FPAGAN	<b>22.3984</b>	<b>0.7418</b>

表 2 GAN 和 FPAGAN 在测试集 B 上的实验结果

Table 2 Experimental results obtained by GAN and FPAGAN on test set B

Method	UIQM	IE	NIQE
GAN	2.9086	4.6889	42.5059
FPAGAN	<b>2.9457</b>	<b>4.6925</b>	<b>37.6927</b>

通过表 2 中的实验结果能够发现,加入 FPA 模块后,增强后的水下图像在非参考图像质量评价指标上的表现更为突出。

#### 4.2.2 在测试集 A 上的对比实验

为验证本文算法的有效性,在测试集 A 上分别进行多次实验,将本文算法与现有的水下图像增强方法(包括 GC,UDCP<sup>[8]</sup>,LDCP<sup>[10]</sup>,UWCNN<sup>[12]</sup>,



FUnIE-GAN<sup>[17]</sup>)进行定性和定量的对比分析。不同方法在测试集 A 上的定性对比分析结果如图 6 所示。可以看出,基于非物理模型的 GC 方法对图像的增强不足,导致图像颜色较浅且出现雾化。而基于物理模型的方法(UDCP,LDCP)会引入意想不到的颜色,图像呈现出红色或蓝色色偏。Li 等<sup>[12]</sup>设计了轻量级的卷积神经网络模型来增强水下图像,有效地去除了雾化,提高了清晰度,但引入了淡黄色

的颜色偏差。Islam 等<sup>[17]</sup>利用配对和非配对的训练集对模型进行训练,从而进行水下图像的增强,该方法通常可以对图像的颜色失真进行有效校正,但偶尔会引入色偏,校正效果不佳。本文算法不仅可以有效地校正颜色偏差,增强对比度,还进一步提升了图像的视觉感知效果。相比之下,本文算法增强的图像在各方面最接近于参考图像,使得增强后的水下图像有更加自然的视觉感知效果。

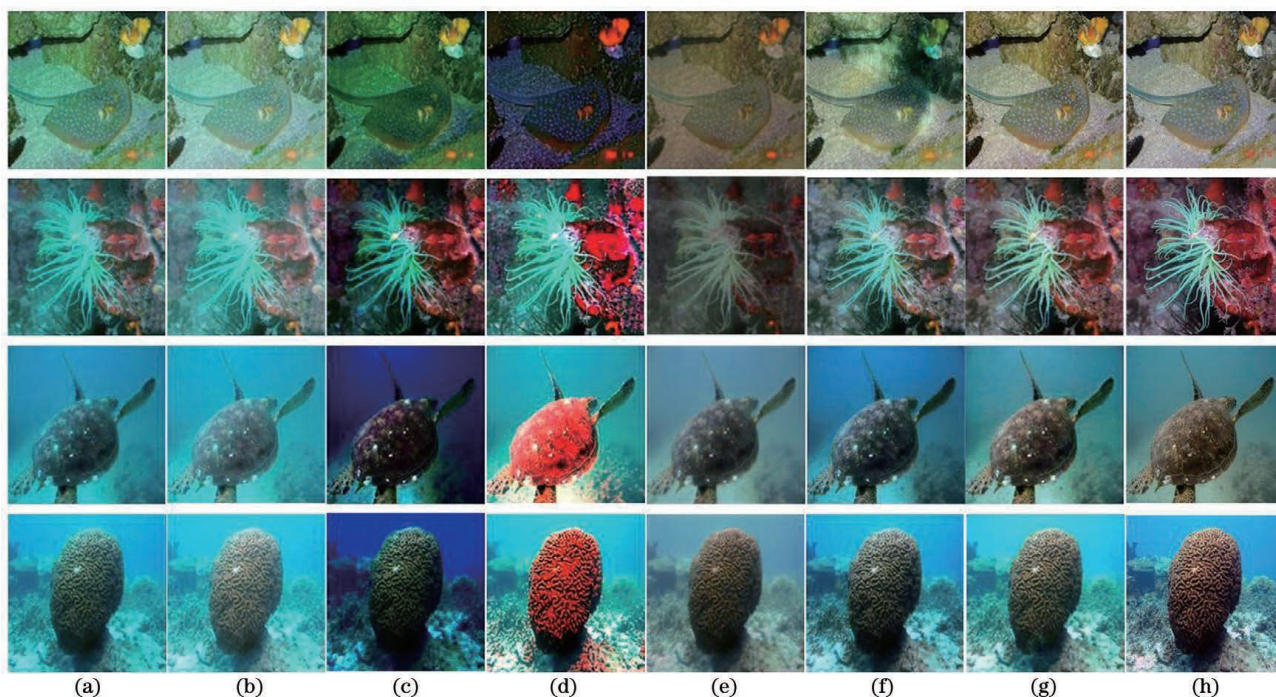


图 6 不同方法在测试集 A 上的定性对比。(a)水下图像;(b)GC;(c)UDCP;(d)LDCP;(e)UWCNN;(f)FUnIE-GAN;(g)本文方法;(h)参考

Fig. 6 Qualitative comparison of different methods on test set A. (a) Underwater images; (b) GC; (c) UDCP; (d) LDCP; (e) UWCNN; (f) FUnIE-GAN; (g) proposed method; (h) reference

为进一步验证本文算法的性能,采用两种全参考图像质量评价指标——PSNR和SSIM,与上述方法进行了定量对比分析。不同方法的定量分析结果如表 3 所示。从表 3 中可以看出,本

文算法在 PSNR 和 SSIM 评价指标上均优于其他对比算法,这表明本文算法能有效恢复图像中的信息、提高对比度,使获得的图像更加真实自然。

表 3 不同方法在测试集 A 上的定量对比

Table 3 Quantitative comparison of different methods on test set A

Metrics	GC	UDCP	LDCP	UWCNN	FUnIE-GAN	Ours
PSNR	15.1830	13.2231	13.9795	16.1344	19.3574	<b>22.3984</b>
SSIM	0.6495	0.5354	0.5422	0.6061	0.6923	<b>0.7418</b>

#### 4.2.3 在测试集 B 上的对比实验

为了更加准确地评估本文算法的性能,将本文算法与上述水下图像增强方法在测试集 B 上分别进行定性和定量对比分析。不同方法在测试集 B 上的定性对比分析如图 7 所示。如图 7(b)所示,

GC 方法虽对色彩偏差进行了校正,但还未完全消除水下图像的蓝绿色调,且带来了雾化效果。从图 7(c)、(d)可以看出,基于物理模型的方法对水下图像进行处理后出现曝光,且引入了红色和蓝色色偏。如图 7(e)所示,文献[12]中的方法在图像的亮

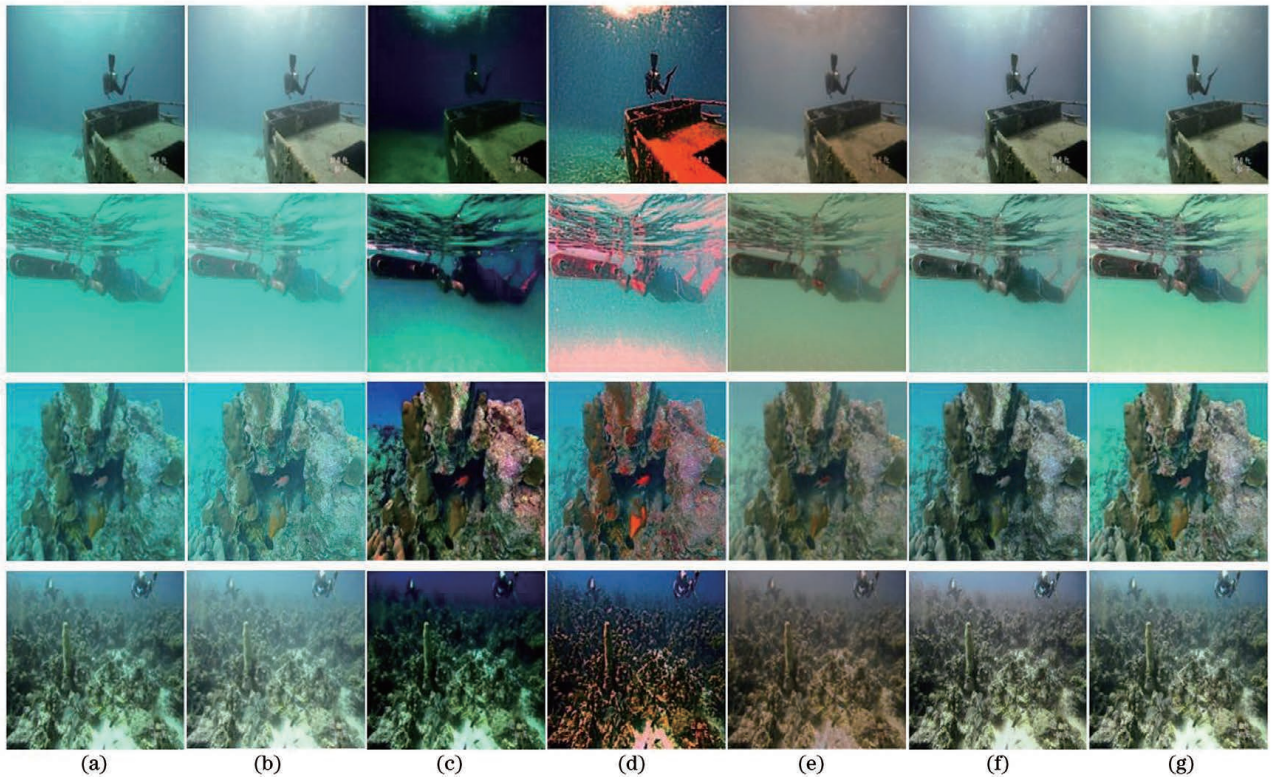


图 7 不同方法在测试集 B 上的定性对比。(a)水下图像;(b)GC;(c)UDCP;(d)LDCP;(e)UWCNN;(f)FUnIE-GAN;  
(g)本文方法

Fig. 7 Qualitative comparison of different methods on test set B. (a) Underwater images; (b) GC; (c) UDCP; (d) LDCP; (e) UWCNN; (f) FUnIE-GAN; (g) proposed method

度和雾化去除等方面还有待改善。如图 7(f) 和图 7(g)所示,文献[17]中的方法和本文算法在图像的蓝绿色调校正、清晰度提升和视觉感知等方面均表现较好。但相比之下,本文算法在颜色校正、去雾和视觉感知等方面的表现更为突出,获得的图像更加清晰自然。因此,本文算法有较好的泛化性能。

为了客观地验证本文算法的性能,选取 4 种非参考图像质量评价指标——UIQM<sup>[26]</sup>、IE、NIQE<sup>[27]</sup>和基于梯度关联的图像质量评估(OG-IQA)<sup>[28]</sup>,对上述不同方法增强后的水下图像实现评估并进行对比

分析。UIQM 是水下图像色度度量(UICM)、水下图像清晰度度量(UISM)和水下图像对比度度量(UIConM)的线性组合,其值越大,图像的视觉质量越好。IE 是衡量图像信息丰富程度的一个重要指标,其值越大,图像所含的信息量越丰富。NIQE 仅利用在自然图像中观察到的信息对图像质量进行评估,其值越小,图像越自然。OG-IQA 利用相邻梯度间的关联性对图像的失真程度进行描述,其值越小,图像质量越好。不同方法在各评价指标上的平均得分如表 4 所示。

表 4 不同方法在测试集 B 上的定量对比

Table 4 Quantitative comparison of different methods on test set B

Metrics	GC	UDCP	LDCP	UWCNN	FUnIE-GAN	Ours
UIQM	2.3361	1.6766	2.0934	2.2210	2.3418	<b>2.9457</b>
IE	4.2288	4.4678	4.5452	4.3699	4.6504	<b>4.6925</b>
NIQE	28.6067	31.0210	<b>27.7121</b>	39.8560	41.1683	37.6927
OG-IQA	-0.4489	-0.6760	-0.7238	-0.5479	-0.7475	<b>-0.8001</b>

通过表 4 可以看出,除 NIQE 评价指标,本文算法在其他三个评价指标上的取值均优于对比算法。

例如,本文算法的 UIQM 评价指标高于对比算法,取值达到 2.9457,说明本文算法可以有效地提升图



像的对比度和清晰度,并可校正颜色偏差。在 IE 和 OG-IQA 评价指标上,本文算法的取值分别为 4.6925 和 -0.8001,说明通过本文算法增强后的图像包含的信息量比较丰富,失真程度较小,视觉质量较高。

## 5 结 论

水下图像增强在水下视觉研究领域具有重要作用。针对水下图像存在模糊、对比度低和颜色偏差等问题,提出了一种基于金字塔注意力机制和生成对抗网络的水下图像增强算法,并根据现有的数据集构造了一个用于训练 FPAGAN 模型的数据集 EUID。FPAGAN 由一个生成网络和一个判别网络构成,并且在生成网络中添加了特征金字塔注意力模块以提升水下图像的增强效果。生成网络用于生成清晰的水下图像,判别网络主要用来促进生成网络生成视觉感知上与参考图像相似的图像。先将原始图像缩放小到  $256 \times 256 \times 3$  大小,再利用生成网络学习原始图像与参考图像的映射关系,并采用多项损失函数,约束生成网络和判别网络进行对抗训练,从而得到 FPAGAN 增强模型。实验结果表明,与对比算法相比,FPAGAN 在 SSIM、UIQM 和 IE 指标上的平均值分别为 0.7418、2.9457 和 4.6925,分别比第二名提升 7.2%、25.8% 和 0.9%。这表明所提算法能有效提高图像清晰度,增强对比度,校正图像颜色,提升图像视觉感知效果。未来的工作将继续优化模型,以进一步提高模型的泛化性能。

## 参 考 文 献

- [1] Shkurti F, Xu A Q, Meghjani M, et al. Multi-domain monitoring of marine environments using a heterogeneous robot team [C] // 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 7-12, 2012, Vilamoura-Algarve, Portugal. New York: IEEE Press, 2012: 1747-1753.
- [2] Hou G J. Research on underwater image enhancement and object recognition algorithms [D]. Qingdao: Ocean University of China, 2015: 19-21.  
侯国家. 水下图像增强与目标识别算法研究[D]. 青岛: 中国海洋大学, 2015: 19-21.
- [3] Jin W P, Guo J C, Qi Q. Underwater image enhancement based on conditional generative adversarial network [J]. Laser & Optoelectronics Progress, 2020, 57(14): 141002.  
晋玮佩, 郭继昌, 祁清. 基于条件生成对抗网络的水下图像增强 [J]. 激光与光电子学进展, 2020, 57(14): 141002.
- [4] Güraksin G E, Köse U, Deperlioglu Ö. Underwater image enhancement based on contrast adjustment via differential evolution algorithm [C] // 2016 International Symposium on INnovations in Intelligent Systems and Applications (INISTA), August 2-5, 2016, Sinaia, Romania. New York: IEEE Press, 2016: 1-5.
- [5] Iqbal K, Odetayo M, James A, et al. Enhancing the low quality images using Unsupervised Colour Correction Method [C] // 2010 IEEE International Conference on Systems, Man and Cybernetics, October 10-13, 2010, Istanbul, Turkey. New York: IEEE Press, 2010: 1703-1709.
- [6] Dai C G, Lin M X, Wang Z, et al. Color compensation based on bright channel and fusion for underwater image enhancement [J]. Acta Optica Sinica, 2018, 38(11): 1110003.  
代成刚, 林明星, 王震, 等. 基于亮通道色彩补偿与融合的水下图像增强 [J]. 光学学报, 2018, 38(11): 1110003.
- [7] Zou P Y, Zhang W D, Shi J Y, et al. Underwater image enhancement algorithm based on fusion of high and low frequency components [J]. Laser & Optoelectronics Progress, 2020, 57(16): 161010.  
邹沛煜, 张卫东, 史金余, 等. 基于高低频分量融合的水下图像增强算法 [J]. 激光与光电子学进展, 2020, 57(16): 161010.
- [8] Drews P, Jr, do Nascimento E, Moraes F, et al. Transmission estimation in underwater single images [C] // 2013 IEEE International Conference on Computer Vision Workshops, December 2-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2013: 825-830.
- [9] Chiang J Y, Chen Y C. Underwater image enhancement by wavelength compensation and dehazing [J]. IEEE Transactions on Image Processing, 2012, 21(4): 1756-1769.
- [10] Yang H Y, Chen P Y, Huang C C, et al. Low complexity underwater image enhancement based on dark channel prior [C] // 2011 Second International Conference on Innovations in Bio-inspired Computing and Applications, December 16-18, 2011, Shenzhen, China. New York: IEEE Press, 2011: 17-20.
- [11] Cai C D, Huo G Y, Zhou Y, et al. Underwater image restoration method based on scene depth estimation and white balance [J]. Laser & Optoelectronics Progress, 2019, 56(3): 031008.  
蔡晨东, 霍冠英, 周妍, 等. 基于场景深度估计和白平衡的水下图像复原 [J]. 激光与光电子学进展,

- 2019, 56(3): 031008.
- [12] Li C Y, Anwar S, Porikli F. Underwater scene prior inspired deep underwater image and video enhancement [J]. *Pattern Recognition*, 2020, 98: 107038.
- [13] 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*, 2020, 29: 4376-4389.
- [14] Fabbri C, Islam M J, Sattar J. Enhancing underwater imagery using generative adversarial networks [C] // 2018 IEEE International Conference on Robotics and Automation (ICRA), May 21-25, 2018, Brisbane, QLD, Australia. New York: IEEE Press, 2018: 7159-7165.
- [15] Yu X L, Qu Y Y, Hong M. Underwater-GAN: underwater image restoration via conditional generative adversarial network [M] // Zhang Z X, Suter D, Tian Y L, et al. *Pattern recognition and information forensics. Lecture notes in computer science*. Cham: Springer, 2019, 11188: 66-75.
- [16] Liu P, Wang G Y, Qi H, et al. Underwater image enhancement with a deep residual framework [J]. *IEEE Access*, 2019, 7: 94614-94629.
- [17] 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.
- [18] Li H C, Xiong P F, An J, et al. Pyramid attention network for semantic segmentation [EB/OL]. (2018-11-25) [2020-10-10]. <https://arxiv.org/abs/1805.10180v1>.
- [19] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [M] // Navab N, Hornegger J, Wells W M, et al. *Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science*. Cham: Springer, 2015, 9351: 234-241.
- [20] Li C, Wand M. Precomputed real-time texture synthesis with Markovian generative adversarial networks [M] // Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9907: 702-716.
- [21] Isola P, Zhu J Y, Zhou T H, et al. Image-to-image translation with conditional adversarial networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5967-5976.
- [22] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution [M] // Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9906: 694-711.
- [23] Ignatov A, Kobyshev N, Timofte R, et al. DSLR-quality photos on mobile devices with deep convolutional networks [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 3297-3305.
- [24] Chen Y S, Wang Y C, Kao M H, et al. Deep photo enhancer: unpaired learning for image enhancement from photographs with GANs [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 6306-6314.
- [25] Horé A, Ziou D. Image quality metrics: PSNR vs. SSIM [C] // 2010 20th International Conference on Pattern Recognition, August 23-26, 2010, Istanbul, Turkey. New York: IEEE Press, 2010: 2366-2369.
- [26] 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.
- [27] Mittal A, Soundararajan R, Bovik A C. Making a “completely blind” image quality analyzer [J]. *IEEE Signal Processing Letters*, 2013, 20(3): 209-212.
- [28] Liu L X, Hua Y, Zhao Q J, et al. Blind image quality assessment by relative gradient statistics and adaboosting neural network [J]. *Signal Processing: Image Communication*, 2016, 40: 1-15.