

# 基于改进CycleGAN的浑浊水体图像增强算法研究

李宝奇 黄海宁\* 刘纪元 刘正君 韦琳哲

(中国科学院声学研究所 北京 100190)

(中国科学院先进水下信息技术重点实验室 北京 100190)

**摘要:** 针对循环生成对抗网络(Cycle Generative Adversarial Networks, CycleGAN)在浑浊水体图像增强中存在质量差和速度慢的问题, 该文提出一种可扩展、可选择和轻量化的特征提取单元BSDK (Bottleneck Selective Dilated Kernel), 并利用BSDK设计了一个新的生成器网络BSDKNet。与此同时, 提出一种多尺度损失函数MLF(Multi-scale Loss Function)。在自建的浑浊水体图像增强数据集TC(Turbid and Clear)上, 该文BM-CycleGAN比原始CycleGAN的精度提升3.27%, 生成器网络参数降低4.15MB, 运算时间减少0.107s。实验结果表明BM-CycleGAN适合浑浊水体图像增强任务。

**关键词:** 图像增强; 生成对抗网络; 循环生成对抗网络; 深度可分离空洞卷积; 多尺度结构相似性

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2022)07-2504-08

DOI: 10.11999/JEIT210400

## Turbid Water Image Enhancement Algorithm Based on Improved CycleGAN

LI Baoqi HUANG Haining LIU Jiyuan LIU Zhengjun WEI Linzhe

(Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China)

(Key Laboratory of Science and Technology on Advanced Underwater Acoustic Signal Processing, Chinese Academy of Sciences, Beijing 100190, China)

**Abstract:** In order to solve the problem of poor quality and slow speed in turbid water image enhancement based on Cycle Generative Adversarial Networks (CycleGAN), a scalable, selective and efficient block Bottleneck Selective Dilated Kernel (BSDK) is proposed, and a new generator network BSDKNet is redesigned by stacking BSDK. At the same time, Multi-scale Loss Function (MLF) is proposed to improve the structural similarity of the clear water image and the generated clear water image. On our turbid water image enhancement dataset Turbid and Clear (TC), the classification accuracy of the proposed BM-CycleGAN is 3.27% higher than that of classical CycleGAN. The generator parameters of BM-CycleGAN is 4.15 MB lower than that of CycleGAN, and the time consuming of BM-CycleGAN is 0.107 s less than that of CycleGAN. The experimental results show that BM-CycleGAN is suitable for turbid water image enhancement.

**Key words:** Image enhancement; Generative adversarial networks; Cycle generative adversarial networks; Selective dilated kernel networks; Multi scale structural similarity

### 1 引言

我国正处于从近海防御到远海防御的过渡期,

海洋装备体系正朝着无人化、智能化的方向快速发展。光学成像具有信息丰富、分辨率高和成本低等优点, 是水下近距离环境感知的主要手段, 广泛用于水下目标确认、水下地形勘测、海底物种监测、海底管道探伤、海底光缆铺设、无人潜航器导航、定位、避障和目标识别等多种场合<sup>[1]</sup>。然而, 与空气介质相比, 水体对于光的衰减作用更为严重, 主要表现在对于光的选择性吸收和强烈散射, 尤其是在浑浊水体条件下, 受多重散射的作用水下光学图像的可视性会变得更差, 增加了水下光学图像的解释难度。

图像增强通过一定技术手段有选择地突出图像

收稿日期: 2021-05-08; 改回日期: 2022-03-01; 网络出版: 2022-03-10

\*通信作者: 黄海宁 hhn@mail.ioa.ac.cn

基金项目: 国家自然科学基金(11904386), 国家基础科研计划重大项目(JCKY2016206A003), 中国科学院青年创新促进会(2019023)

Foundation Items: The National Natural Science Foundation of China (11904386), State Administration of Science, Technology and Industry for National Defence (JCKY2016206A003), Youth Innovation Promotion Association of Chinese Academy of Sciences (2019023)

中感兴趣的特征或者抑制图像中某些不需要的特征,使图像与视觉响应特性相匹配。Hitam等人<sup>[2]</sup>针对水下图像增强而提出了一种混合对比度受限自适应直方图均衡(Contrast Limited Adaptive Histogram Equalization, CLAHE)颜色模型。该方法通过在RGB和HSV颜色模型上执行CLAHE算法,再使用欧几里德范数将两组结果结合在一起,来降低水下图像噪声和伪影。Ghani等人<sup>[3]</sup>提出一种双图像瑞利拉伸对比度受限自适应直方图的新方法,该方法结合了全局和局部对比度校正,通过增加图像细节改善水下图像的可见度。Li等人<sup>[4]</sup>针对水下光线吸收和散射的影响导致的目标偏色与模糊问题,提出了一种基于颜色校正和水下图像去雾的混合方法,以改善退化的水下图像的视觉质量。Deng等人<sup>[5]</sup>提出了一种水下图像增强算法RLSCD (Removing Light Source Color and Dehazing), RLSCD有效改善水下图像偏色问题。虽然上述图像增强算法可以实现一定浑浊度范围下的图像恢复,但是恢复后的图像与低浑浊度下图像存在明显区别,无法恢复到同等画质。

深度学习<sup>[6-8]</sup>的出现从根本上改变了人们处理复杂和不确定性问题的方法,与传统的机器学习和信号处理方法相比,深度学习模拟人类视觉神经系统的层次体系,含有更多的隐含单元层,通过对原始数据逐层的非线性变换,可以得到更高层次的、更加抽象的特征表达,高层次的表达能够强化输入数据的区分能力,同时削弱不相关因素的不利影响<sup>[9,10]</sup>。Goodfellow等人<sup>[11]</sup>提出生成对抗网络(Generative Adversarial Networks, GAN)。GAN由生成式模型和判别式模型组成,生成模型用于捕捉样本数据的分布,判别模型用于判断生成数据是否正确,GAN为获取大量的样本数据提供了非常好的解决思路。为了提高GAN模型的数据生成质量和稳定性,Radford等人<sup>[12]</sup>提出了深度卷积生成对抗网络(Deep Convolution Generative Adversarial Networks, DCGAN)。Arjovsky等人<sup>[13]</sup>提出WGAN (Wasserstein GAN)模型。WGAN使用Earth-Mover保证了生成样本的多样性。随着GAN理论体系的完善<sup>[14,15]</sup>,其研究成果促进了图像增强的发展。Isola等人<sup>[16]</sup>提出了Pix2Pix模型。Pix2Pix实现了基于GAN的图像迁移任务,但Pix2Pix要求的输入必须是成对图像。为了解决输入图像成对的限制,Zhu等人<sup>[17]</sup>基于GAN和对偶学习设计了循环一致性对抗网络(CycleGAN)。CycleGAN利用双重任务之间的反馈信号对模型进行训练来实现非配对图像迁移学习任务。在此基础上,Fabbri等人<sup>[18]</sup>提出了一

种基于CycleGAN的水下图像颜色校正方法,降低水下图像的偏色和噪声的影响。不过,CycleGAN生成器网络特征提取层采用固定大小的卷积核容易造成图像信息的丢失;另外,原始循环一致损失函数仅是对输入图像和重构图像的颜色约束,无法满足生成图像与真实图像结构上的相似。对生成器网络而言,通常的做法是增加模型的深度<sup>[19]</sup>或者增加模型的尺度<sup>[20]</sup>。随着对卷积神经网络研究的深入,Li等人<sup>[21]</sup>提出了SKNet网络模块。SKNet在多尺度模块的基础上设计了一个特征选择机制,即通过学习能保留最有价值的多尺度特征层,SKNet采用大小不同卷积核,显然会增加生成器网络的参数和计算量。为了保证生成图像与输入图像的结构相似性,Huang等人<sup>[22]</sup>提出了基于SSIM (Structural Similarity index)的循环生成对抗网络SS-GycleGAN。SS-GycleGAN的引入会增加输入图像与重构图像之间结构的相似性,进而提高生成图像的质量。不过,SS-GycleGAN仅在一个尺度上对输入图像和重构图像进行约束。

针对上述方法存在的问题,本文首先提出一种兼顾多尺度通道可选择和计算量低的特征提取单元BSDK (Bottleneck Selective Dilated Kernel),并利用BSDK设计了一个新的生成器网络BSDKNet。接着,提出了一种基于多尺度结构相似性的循环一致损失函数MLF (Multi-scale Loss Function),从多个尺度上保证输入图像(浑浊水体图像)与生成图像(清澈水体图像)结构信息相似。最后,在CycleGAN框架内通过结合BSDKNet和MLF(BM-CycleGAN)实现浑浊水下目标图像到清澈水下目标图像快速、准确的迁移生成。

## 2 基于改进CycleGAN的浑浊水体图像增强算法

本文通过改进生成器网络和循环一致损失函数来改善浑浊水体图像增强的精度和速度。本节首先介绍浑浊水体图像增强算法BM-CycleGAN网络结构、多尺度损失函数MLF和训练过程。接着介绍BSDK特征提取单元和BSDKNet生成器网络结构。

### 2.1 BM-CycleGAN

BM-CycleGAN由一组对称的生成器 $\{G, F\}$ 、两个判别器 $\{D_X, D_Y\}$ 和循环一致损失函数MLF组成,网络结构如图1所示。在BM-CycleGAN上路,浑浊图像 $x$ 经生成器 $G$ 生成清澈生成图像 $G(x)$ ,再经生成器 $F$ 生成浑浊重构图像 $F(G(x))$ 。在BM-CycleGAN下路,清澈图像 $y$ 经生成器 $F$ 生成浑浊生成图像 $F(y)$ ,再经生成器 $G$ 生成清澈重构图像 $G(F(y))$ 。生成器 $\{G, F\}$ 的网络结构一样、将在2.2节做重点

介绍。在BM-CycleGAN中路, 判别器 $D_X$ 负责区分浑浊图像 $x$ 和浑浊生成图像 $F(y)$ 的真假; 判别器 $D_Y$ 负责区分清晰图像 $y$ 和清晰生成图像 $G(x)$ 的真假。判别器 $\{D_X, D_Y\}$ 的网络结构一样, 均采用PatchGAN结构<sup>[17]</sup>。在BM-CycleGAN上路和下路的外侧, 循环一致损失函数MLF通过减少输入图像与重构图像的差异来引导模型的训练。

假定 $x \sim p_{\text{data}}(x)$ ,  $y \sim p_{\text{data}}(y)$ 代表 $X, Y$ 两个域中数据服从的概率分布。BM-CycleGAN生成器 $G: X \rightarrow Y$ 的单项损失函数为

$$L_{\text{GAN}}(G, D_Y, X, Y) = E_{y \sim p_{\text{data}}(y)} [\lg D_Y(y)] + E_{x \sim p_{\text{data}}(x)} [1 - D_Y(G(x))] \quad (1)$$

其中,  $G$ 会生成类似于 $Y$ 中的清晰样本 $G(x)$ ,  $D_Y$ 的任务是区分 $G(x)$ 和真实样本 $y$ 。同样生成器 $F: Y \rightarrow X$ 的单向损失函数为

$$L_{\text{GAN}}(F, D_X, Y, X) = E_{y \sim p_{\text{data}}(y)} [\lg D_X(x)] + E_{x \sim p_{\text{data}}(x)} [1 - D_X(F(x))] \quad (2)$$

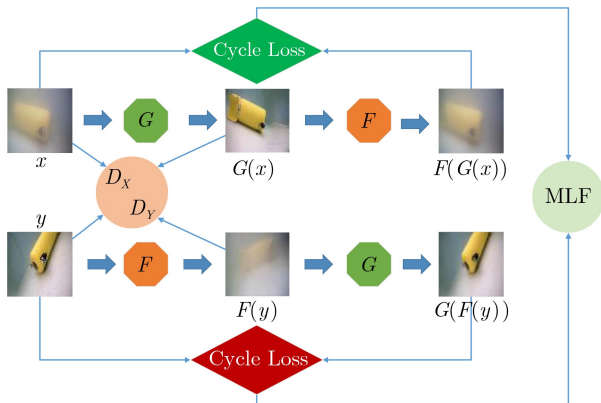


图1 BM-CycleGAN网络结构

其中,  $F$ 会生成类似于 $X$ 中浑浊样本 $F(y)$ ,  $D_X$ 的任务是区分 $F(y)$ 和真实样本 $x$ 。

多尺度循环一致损失函数MLF定义为

$$L_{\text{MLF}}(G, F) = (10 - \lambda) \times L_1(G, F) + \lambda \times L_{\text{MS-SSIM}}(G, F) \quad (3)$$

$$L_1(G, F) = E_{X \sim p_{\text{data}}(x)} [\|F(G(x)) - x\|_1] + E_{Y \sim p_{\text{data}}(y)} [\|G(F(y)) - y\|_1] \quad (4)$$

$$L_{\text{MS-SSIM}}(G, F) = E_{X \sim p_{\text{data}}(x)} [1 - \text{MS-SSIM}(F(G(x)), x)] + E_{Y \sim p_{\text{data}}(y)} [1 - \text{MS-SSIM}(G(F(y)), y)] \quad (5)$$

其中,  $L_1(G, F)$ 为施加 $F(G(x)) - x$ 和 $G(F(y)) - y$ 的 $L_1$ 范数的数学期望约束, 它负责学习图像的颜色特征;  $L_{\text{MS-SSIM}}(G, F)$ 为施加 $F(G(x)) - x$ 和 $G(F(y)) - y$ 的多尺度结构相似性函数的数学期望, 它负责学习图像的结构特征。本文权重系数 $\lambda=6$ 。

将式(1)、式(2)与多尺度循环一致损失函数MLF合并可以得到BM-CycleGAN总损失函数为

$$L(G, F, D_X, D_Y) = L_{\text{GAN}}(G, D_Y, X, Y) + L_{\text{GAN}}(F, D_X, Y, X) + L_{\text{MLF}}(G, F) \quad (6)$$

## 2.2 BSDK模块和BSDKNet生成器网络

BBB(Bottleneck Building Block)模块、SK模块、BSK模块和BSDK模块的关系如图2所示。图2(a)为CycleGAN采用的BBB模块, 它由两个 $1 \times 1$ 的卷积层和1个 $3 \times 3$ 的卷积层组成, 第1个 $1 \times 1$ 卷积层用于通道压缩, 第2个 $1 \times 1$ 卷积层用于通道扩张。图2(b)为SK模块, 它由分裂层、多尺度卷积层、融合层和选择层组成。输入特征经分裂层送入多尺度分组卷积层, 得到多个尺度的输出特征; 接着, 在融合层对输出特征进行相加融合; 最

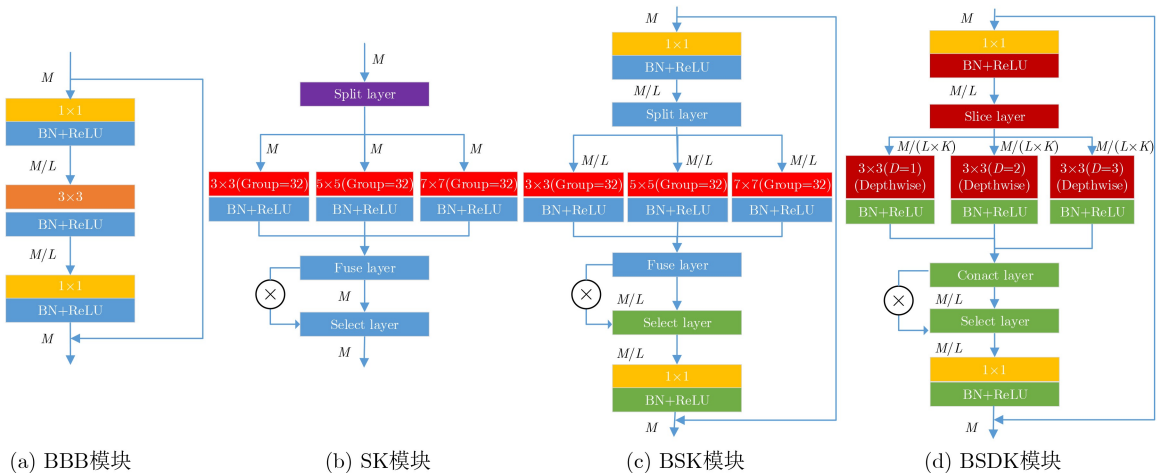


图2 BSDK特征提取单元



后,在选择层对多尺度特征进行筛选来提高提取特征的准确性。图2(c)是BSK模块,是SK模块的BBB形式。输入特征首先经过第1个 $1 \times 1$ 的卷积层进行通道压缩,然后经过SK模块进行多尺度特征选择,最后经过第2个 $1 \times 1$ 的卷积层进行通道扩张。图2(d)为BSDK模块,它在图2(c)结构的基础上将分裂层替换成分割层,分割层实现输入特征通道等分;另外,将大尺寸卷积核( $5 \times 5$ 和 $7 \times 7$ )替换为深度可分离空洞卷积<sup>[23]</sup>,深度可分离空洞卷积首先将标准卷积分解成深度可分离卷积和点卷积<sup>[24]</sup>,然后在深度可分离卷积中引入1个称作空洞率<sup>[25,26]</sup>的新参数,并利用扩张率控制卷积核处理数据时各值的间距;最后,利用拼接层实现多尺度特征通道拼接融合。

对于一个任意的输入特征 $T \in \Phi^{H \times W \times M}$ ,其中 $H \times W$ 为输入特征的尺寸, $M$ 为输入特征的通道数。输入特征 $T$ 进入BSDK模块的两个支路网络:左侧支路负责多尺度特征提取和选择;右侧支路保持输入特征 $T$ 不变,并最后与左侧支路网络的输出特征相加。对于左侧支路网络,输入特征 $T$ 首先经过压缩层,其输出特征的数学表达式为

$$T \rightarrow U \in \Phi^{H \times W \times (M/L)} \quad (7)$$

其中, $T$ 为原始输入特征, $U$ 为经过压缩层后的特征,压缩层的卷积核尺寸为 $1 \times 1$ ,卷积核的数量为输入特征通道的 $L$ 分之一,即 $M/L$ 。

随后,输出特征 $U$ 经分割层送入多尺度深度可分离空洞卷积层,其输出特征的数学表达式为

$$U_k \rightarrow V_k \in \Phi^{H \times W \times (M/(L \times K))}, k=1, 2, \dots, K \quad (8)$$

其中, $V_k$ 为深度可分离空洞卷积层输出的特征图,特征图尺寸为 $H \times W$ ,通道数为 $M/(L \times K)$ , $K$ 为多尺度实际空洞滤波器(Practical Dilated Filter, RDF)的类型数,例如RDF为3, 5和7,则 $K=3$ 。RDF尺寸与空洞率之间的关系为

$$R_{\text{rdf}} = R + (\kappa - 1)(R - 1) \quad (9)$$

其中, $R_{\text{rdf}}$ 为该层RDF尺寸, $R$ 为该层卷积核尺寸, $\kappa$ 为该层空洞率大小。例如,一个卷积核尺寸为 $3 \times 3$ ,空洞率2的空洞卷积层,RDF的实际覆盖范围为 $R_{\text{rdf}}=5$ 。增大空洞率 $\kappa$ 可以扩大卷积层的感受野。因此,BSDK可以用更少的参数和计算量实现与BSK相当的特征提取能力。

接着,对 $K$ 个多尺度深度可分离空洞卷积层的输出 $V_k$ 在通道项进行拼接融合,其输出特征的数学表达式为

$$V_k \rightarrow V \in \Phi^{H \times W \times (M/L)}, k=1, 2, \dots, K \quad (10)$$

其中, $V$ 为拼接融合后的输出特征,输出特征图的尺寸 $H \times W$ ,通道数为 $M/L$ 与压缩后的通道数相等。

然后,对拼接融合后的输出特征 $V$ 进行通道选择,多尺度通道选择系数的数学表达式为

$$s = \text{softmax}(f_{\text{cs}}(f_c(P_g(V)))) \quad (11)$$

其中, $s$ 为多尺度通道的选择系数, $s \in \Phi^{1 \times (M/L)}$ ;  $P_g(\cdot)$ 为全局池化函数,输出特征维度为 $\Phi^{1 \times (M/L)}$ ;  $f_c$ 为第1全连接层,输出特征维度为 $\Phi^{1 \times d}$ ,其中 $d=32$ ;  $f_{\text{cs}}$ 为第2全连接层,输出特征维度为 $\Phi^{1 \times (M/L)}$ ;  $\text{softmax}(\cdot)$ 为归一化指数函数。多尺度通道选择系数(式(5))与输出特征相乘得到选择后的输出特征,其数学表达式为

$$V' = s \cdot V \quad (12)$$

其中, $V'$ 为通道选择后的多尺度通道特征。

接着,对 $V'$ 进行通道扩张,数学表达式为

$$V' \rightarrow T' \in \Phi^{H \times W \times M} \quad (13)$$

其中, $T'$ 为通道压缩后的多尺度通道特征。

通过上面的计算,最后可以得到ESK模块的输出特征数学表达式为

$$O = T + T' \quad (14)$$

其中, $O$ 为ESK模块的输出特征, $O \in \Phi^{H \times W \times M}$ ,特征图尺寸为 $H \times W$ ,通道数为 $M$ 。

BSDKNet沿用CycleGAN原始生成器的网络架构,包括编码器、迁移器和解码器组成,其网络结构如图3所示。编码器由保持层(第1卷积层)和下采样层(第2~5卷积层)组成,迁移器由BSDK模块(第6~14卷积层)组成,解码器由上采样层(第15~18卷积层)和输出保持层(第19卷积层)组成。本文BSDK卷积核尺寸为 $3 \times 3$ ,设置4个不同空洞率的特征层,空洞率分别为1, 2, 3和4,分别对应SK模块卷积 $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ 和 $9 \times 9$  4个尺度的卷积核。

以BM-CycleGAN上路生成器 $G$ 说明浑浊水体图像在BSDK网络中特征尺寸变化过程。浑浊水体图像首先经输出保持层,尺寸保持不变、通道数由3变为64。接着,经过下采样层,通道数逐层放大1倍、特征尺寸减少1倍,总共4个下采样层,最终特征尺寸变为输入图像的 $1/16$ 、通道数为256。然后,经过9个由BSDK堆叠组成的生成器,每一层的尺寸和通道数保持不变。随后,经过上采样层,通道数逐层减少1倍、尺寸增大1倍,总共4个上采样层,最终输出特征尺寸与输入图像相等、通道数为64。最后,经过输出保持层,生成与输入图像通道数相同的清澈水下目标图像。

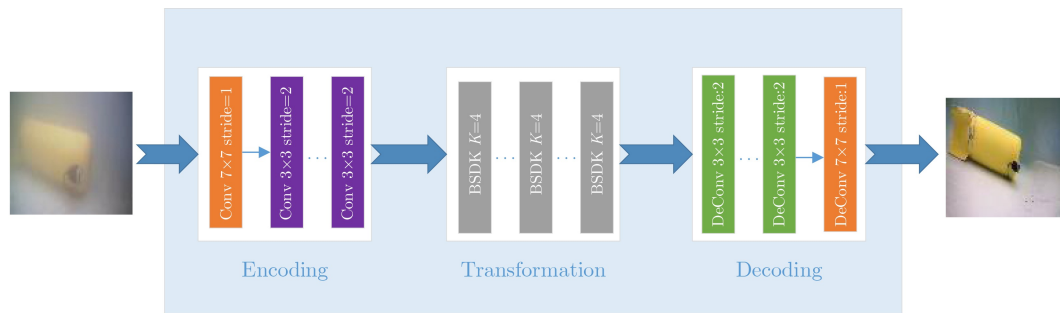


图3 生成器网络BSDKNet结构

### 3 实验分析

为了验证生成器网络BSDKNet和循环一致损失函数MLF对CycleGAN性能的影响,实验以分类准确率、生成器网络参数大小和生成器运算时间作为模型的定量评价指标。设计实验1,以原始CycleGAN模型为参考,比较分析BM-CycleGAN与其它模型之间的性能差异。设计实验2,比较分析BSDKNet网络和MLF损失函数对BM-CycleGAN的影响。设计实验3,比较分析不同权重系数的MLF对BM-CycleGAN性能影响。设计实验4,比较分析不同多尺度系数的BSDKNet对BM-CycleGAN性能影响。

#### 3.1 水下浑浊水体图像增强数据集

为了更好地检验BM-CycleGAN对浑浊水体图像的增强效果,我们搭建了一个浑浊水体和清澈水体环境,其中浑浊水体环境的光衰减系数大于0.6,清澈水体环境的光衰减系数小于0.2。通过水下摄像机对两种水体下的目标进行图像采集,并创建浑浊清澈水体数据集TC(Turbid and Clear)。数据集TC浑浊条件下的圆柱形目标和长方形目标的数量分别是72和63,清澈条件下的圆柱形目标和长方形目标的数量分别是79和65。图像的尺寸为300像素×300像素,剪裁至256像素×256像素送入模型训练和测试。

#### 3.2 实验1: 浑浊水体图像增强算法性能比较

本实验比较分析CycleGAN, SS-CycleGAN, BSK-CycleGAN和BM-CycleGAN之间的性能差异,其中CycleGAN表示原始循环一致生成对抗网络;SS-CycleGAN的生成器网络基于BBB模块,循环一致损失函数采用文献[22];BSK-CycleGAN的生成器网络基于BSK模块,循环一致损失函数采用文献[17];BM-CycleGAN的生成器网络为BSDKNet,循环一致损失函数为MLF。保存CycleGAN, SS-CycleGAN, BSK-CycleGAN和BM-CycleGAN迭代1000次时的生成模型,并生成清澈水下目标图像。记录4个模型分类准确率、参数大小和运算时间,实验结果如表1所示。

表1 浑浊水体图像增强算法性能比较

	分类准确率(%)	参数大小(MB)	运算时间(s)
CycleGAN	95.10	10.97	0.127
SS-CycleGAN	95.92	10.97	0.127
BSK-CycleGAN	96.73	18.21	0.054
BM-CycleGAN	98.37	6.82	0.020

从表1可以发现, BM-CycleGAN 的分类准确率高於CycleGAN, SS-CycleGAN和BSK-CycleGAN, 分别高3.27%, 2.54%和1.64%; 参数比CycleGAN, SS-CycleGAN和BSK-CycleGAN 分别减少4.15MB, 4.15MB和11.39MB; 运算时间比CycleGAN, SS-CycleGAN和BSK-CycleGAN 分别降低0.107 s, 0.107 s和0.034 s。在参数和运算时间方面, CycleGAN和SS-CycleGAN的生成器基于BBB模块(标准卷积), 计算成本高; BSK-CycleGAN的生成器网络基于BSK模块(分组卷积), 计算成本较低, 不过BSK模块每个多尺度特征层的输入特征数量与BBB模块相同, 模型参数较大; BM-CycleGAN的生成器基于BSDK模块, 为深度可分离卷积, 计算成本最低, 同时BSDK多尺度特征层的输入特征数量之和与BBB模块相等, 因此模型参数依然保持较低水平。综合考虑分类准确率、参数大小和运算时间3个评价指标, BM-CycleGAN优于其他模型。

为了更直观地说明本文算法的性能, 将原始圆柱形目标浑浊图像(图4(a))和原始长方形目标浑浊图像(图4(c))分别送入训练好的BM-CycleGAN的生成模型中, 得到清澈圆柱形目标生成图像(图4(b))和清澈长方形目标生成图像(图4(d)), 图4(b)和图4(d)与原始清澈目标图像特性以及背景纹理非常接近, 达到了假乱真的水平。

#### 3.3 实验2: BSDKNet生成器网络和MLF损失函数对BM-CycleGAN的影响

本实验比较BSDKNet生成器网络和MLF损失函数两个因素对BM-CycleGAN的影响。当两个因素都不起作用时, 记做CycleGAN(原始循环一致生

成对抗网络); 当仅MLF损失函数起作用时, 记做M-CycleGAN; 当仅BSDKNet生成器网络起作用时, 记做B-CycleGAN; 当两个因素都起作用时, 记做BM-CycleGAN(本文方法)。保存4种模型迭代1000次时的生成模型, 并生成清澈水下目标图像。记录模型的分类准确率、参数大小和运算时间, 实验结果如表2所示。

从表2可以发现, M-CycleGAN, B-CycleGAN和BM-CycleGAN比CycleGAN的分类准确率分别提高0.99%, 2.17%和3.27%, 其中B-CycleGAN比M-CycleGAN的分类准确率提高1.18%。由此可见, BSDKNet生成器网络和MLF损失函数两个因素均能改善CycleGAN的性能。M-CycleGAN仅启用MLF损失函数, 在MLF的约束下, 分类准确率比CycleGAN提高0.99%, 模型的参数大小和运算时间相同。B-CycleGAN仅启用BSDKNet生成器网络, 在BSDKNet的作用下, 分类准确率比CycleGAN提高2.17%, 模型的参数大小减少4.15MB, 运算时间降低0.107s。综合考虑分类准确率、参数大小和运算时间3个评价指标, BSDKNet生成器网络比MLF损失函数对CycleGAN的改善作用更大。此外, 从表2也可以进一步发现和验证, 模型的参数大小和运算时间由生成器网络决定, 与循环一致损失函数无关。

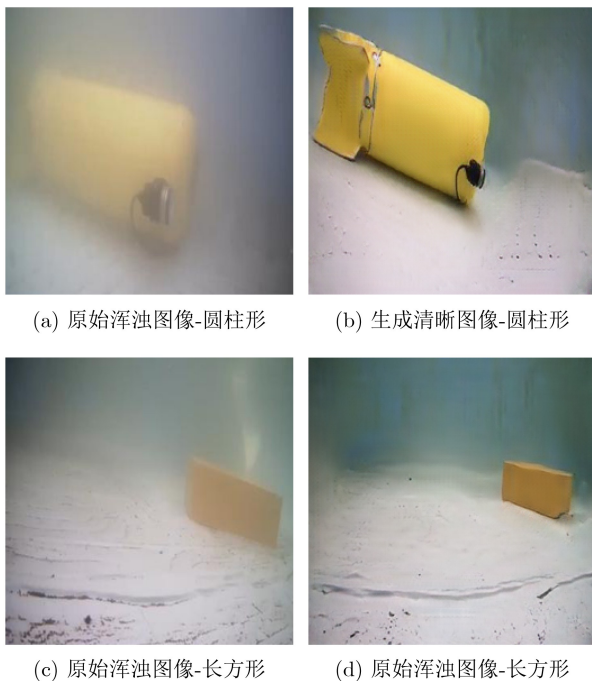


图4 浑浊水体增强效果

### 3.4 实验3: 不同权重系数MLF对BM-CycleGAN的影响

本实验比较不同权重系数MLF对BM-CycleGAN的影响, 生成器为BSDKNet, 通过改变MLF权重系数得到不同条件下BM\_CycleGAN模型, 其中多尺度循环一致损失函数MLF权重系数分别为1, 2, 3, 4, 5, 6, 7, 8和9。保存9种模型迭代1000次时的生成模型, 并生成清澈水下目标图像。由于改变MLF的权重系数不影响生成器的参数大小和模型的运算时间, 本实验仅记录分类准确率, 实验结果如表3所示。

从表3可以发现, BM-CycleGAN的分类准确率呈现先增大后减小的规律, 在权重系数等于6的条件下, BM-CycleGAN的分类准确率最高等于98.37%。也就是说, 合理的权重系数能提高BM-CycleGAN的性能。

### 3.5 实验4: 不同多尺度系数的BSDKNet对BM-CycleGAN性能的影响

本实验比较不同多尺度系数BSDKNet对BM-CycleGAN性能的影响。循环一致损失函数为MLF, 权重系数等于6。通过改变BSDKNet多尺度系数得到不同条件下的BM-CycleGAN模型, 轻量化生成器BSDKNet的多尺度系数分别为1, 2, 4和8, 保存4种模型迭代1000次时的生成模型, 并生成清澈水下目标图像, 记录分类准确率、参数大小和运算时间, 实验结果如表4所示。

从表4可以看出, BM-CycleGAN的分类准确率随 BSDKNet多尺度系数增大而增加, 模型参数保持不变, 运算时间逐渐增大。可见, BSDKNet模块在增大多尺度系数的条件下, 模型的参数并没有增加。计算时间的增加主要是由于算法在计算多尺度卷积时, 采用的是逐个操作。本文兼顾分类准确率和运算时间, 本文多尺度系数等于4。

表2 BSDKNet生成器网络和MLF损失函数对BM-CycleGAN性能的影响

	BSDKNet	MLF	分类准确率(%)	参数大小(MB)	运算时间(s)
CycleGAN			95.10	10.97	0.127
M-CycleGAN		✓	96.09	10.97	0.127
B-CycleGAN	✓		97.27	6.82	0.020
BM-CycleGAN	✓	✓	98.37	6.82	0.020

表3 不同权重系数(1-9)MLF对BM-CycleGAN性能的影响

	1	2	3	4	5	6	7	8	9
分类准确率(%)	95.10	95.51	96.73	97.14	97.55	98.37	97.96	97.14	96.33



表4 不同多尺度系数(1,2,4,8)条件下的BSDKNet对BM-CycleGAN性能的影响

	1	2	4	8
分类准确率(%)	95.10	97.96	98.37	98.78
参数大小(MB)	6.82	6.82	6.82	6.82
运算时间(s)	0.006	0.016	0.021	0.030

### 3.6 讨论

CycleGAN为浑浊水体图像增强提供了解决思路,实验从分类准确率、参数大小和运算时间3个方面比较了本文浑浊水体图像增强算法BM-CycleGAN与经典算法(CycleGAN)和最新算法(SS-CycleGAN和BSK-CycleGAN)性能上的差异,也进一步分析了生成器网络和循环一致损失函数两个因素及其参数选取如何影响BM-CycleGAN的性能。对于CycleGAN而言,优化生成器网络和循环一致损失函数都能改善浑浊水体图像增强效果,不过循环一致损失函数不影响生成器网络运算时间和参数大小,它只影响模型的训练复杂度,实验1和实验2的结果也验证了这一事实。BSDK模块在BSK模型的基础上,利用深度可分离卷积降低了模型参数和计算量,同时改进通融合方式提高了BSDKNet的特征提取能力,在性能上优于BSK模块。MLF通过引入多尺度结构相似性函数,增加了生成图像与原始图像在结构上的相似性,使增强后的图像与原始清澈水体图像更相似。

## 4 结束语

浑浊水体图像增强具有重要的理论研究和实际应用价值。在CycleGAN模型框架内,本文提出了一种新的生成器网络BSDKNet和循环一致损失函数MLF,新的生成器网络和循环损失函数有效提升了CycleGAN对浑浊水体图像增强的精度和效率,经理论分析和仿真实验证明了新方法的有效性。BM-CycleGAN可实现与低浑浊度水下光学图像相同画质的图像增强效果,增强后水下光学图像具有更强的解释性。

对浑浊水体增强任务,下一步的研究重点包括:(1)研究具有先验知识的浑浊水体图像增强模型;(2)研究更多种类的浑浊水体图像增强模型。

### 参考文献

- [1] HMUE P M and PUMRIN S. Image enhancement and quality assessment methods in turbid water: A review article[C]. 2019 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia), Bangkok, Thailand, 2019: 59-63. doi: [10.1109/ICCE-Asia46551.2019.8942216](https://doi.org/10.1109/ICCE-Asia46551.2019.8942216).
- [2] HITAM M S, AWALLUDIN E A, YUSSOF W N J H W, et al. Mixture contrast limited adaptive histogram equalization for underwater image enhancement[C]. 2013 International Conference on Computer Applications Technology (ICCAT), Sousse, Tunisia, 2013: 1-5. doi: [10.1109/ICCAT.2013.6522017](https://doi.org/10.1109/ICCAT.2013.6522017).
- [3] GHANI A S A and ISA N A M. Enhancement of low quality underwater image through integrated global and local contrast correction[J]. *Applied Soft Computing*, 2015, 37: 332-344. doi: [10.1016/j.asoc.2015.08.033](https://doi.org/10.1016/j.asoc.2015.08.033).
- [4] LI Chongyi, GUO Jichang, GUO Chunle, et al. A hybrid method for underwater image correction[J]. *Pattern Recognition Letters*, 2017, 94: 62-67. doi: [10.1016/j.patrec.2017.05.023](https://doi.org/10.1016/j.patrec.2017.05.023).
- [5] DENG Xiangyu, WANG Huigang, and LIU Xing. Underwater image enhancement based on removing light source color and Dehazing[J]. *IEEE Access*, 2019, 7: 114297-114309. doi: [10.1109/ACCESS.2019.2936029](https://doi.org/10.1109/ACCESS.2019.2936029).
- [6] LECUN Y, BENGIO Y, and HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553): 436-444. doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- [7] KWOK R. Deep learning powers a motion-tracking revolution[J]. *Nature*, 2019, 574(7776): 137-138. doi: [10.1038/d41586-019-02942-5](https://doi.org/10.1038/d41586-019-02942-5).
- [8] WANG Shiqiang. Efficient deep learning[J]. *Nature Computational Science*, 2021, 1(3): 181-182. doi: [10.1038/s43588-021-00042-x](https://doi.org/10.1038/s43588-021-00042-x).
- [9] KUANG Wenhuan, YUAN Congcong, and ZHANG Jie. Real-time determination of earthquake focal mechanism via deep learning[J]. *Nature Communications*, 2021, 12(1): 1432. doi: [10.1038/s41467-021-21670-x](https://doi.org/10.1038/s41467-021-21670-x).
- [10] YANG X S. Data Mining and Deep Learning[M]. YANG X S. Nature-Inspired Optimization Algorithms (Second Edition). London: Academic Press, 2021: 239-258.
- [11] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Networks[C]. Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, Canada, 2014: 2672-2680.
- [12] RADFORD A, METZ L, and CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[C]. 4th International Conference on Learning Representations, San Juan, Puerto Rico, 2016.
- [13] ARJOVSKY M, CHINTALA S, and BOTTOU L. Wasserstein GAN[J]. arXiv Preprint arXiv: 1701.07875, 2017.
- [14] CHEN Xi, DUAN Yan, HOUTHOOFT R, et al. InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets[C]. Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, Barcelona, Spain, 2016: 2172-2180.
- [15] XU Qiantong, HUANG Gao, YUAN Yang, et al. An

- empirical study on evaluation metrics of generative adversarial networks[J]. arXiv preprint arXiv: 1806.07755, 2018.
- [16] ISOLA P, ZHU Junyan, ZHOU Tinghui, *et al.* Image-to-image translation with conditional adversarial networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 2017: 5967–5976. doi: [10.1109/CVPR.2017.632](https://doi.org/10.1109/CVPR.2017.632).
- [17] ZHU Junyan, PARK T, ISOLA P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks[C]. 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017: 2242–2251. doi: [10.1109/ICCV.2017.244](https://doi.org/10.1109/ICCV.2017.244).
- [18] FABBRI C, ISLAM M J, and SATTAR J. Enhancing underwater imagery using generative adversarial networks[C]. 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 2018: 7159–7165. doi: [10.1109/ICRA.2018.8460552](https://doi.org/10.1109/ICRA.2018.8460552).
- [19] XIE Saining, GIRSHICK R, DOLLÁR P, *et al.* Aggregated residual transformations for deep neural networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 2017: 5987–5995. doi: [10.1109/CVPR.2017.634](https://doi.org/10.1109/CVPR.2017.634).
- [20] SZEGEDY C, IOFFE S, VANHOUCHE V, *et al.* Inception-v4, inception-ResNet and the impact of residual connections on learning[C]. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, USA, 2017: 4278–4284.
- [21] LI Xiang, WANG Wenhai, HU Xiaolin, *et al.* Selective kernel networks[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, USA, 2019: 510–519. doi: [10.1109/CVPR.2019.00060](https://doi.org/10.1109/CVPR.2019.00060).
- [22] HUANG Xuejun, WEN Liwu, and DING Jinshan. SAR and optical image registration method based on improved CycleGAN[C]. 2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR), Xiamen, China, 2019: 1–6. doi: [10.1109/APSAR46974.2019.9048448](https://doi.org/10.1109/APSAR46974.2019.9048448).
- [23] 李宝奇, 贺昱曜, 张伟, 等. 基于并行附加特征提取网络的SSD地面小目标检测模型[J]. 电子学报, 2020, 48(1): 84–91. doi: [10.3969/j.issn.0372-2112.2020.01.010](https://doi.org/10.3969/j.issn.0372-2112.2020.01.010).
- LI Baoqi, HE Yuyao, QIANG Wei, *et al.* SSD with parallel additional feature extraction network for ground small target detection[J]. *Acta Electronica Sinica*, 2020, 48(1): 84–91. doi: [10.3969/j.issn.0372-2112.2020.01.010](https://doi.org/10.3969/j.issn.0372-2112.2020.01.010).
- [24] HOWARD A G, ZHU Menglong, CHEN Bo, *et al.* Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv Preprint arXiv: 1704.04861, 2017.
- [25] CHEN L C, PAPANDREOU G, KOKKINOS I, *et al.* DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834–848. doi: [10.1109/TPAMI.2017.2699184](https://doi.org/10.1109/TPAMI.2017.2699184).
- [26] QIN Yanjun, LUO Haiyong, ZHAO Fang, *et al.* NDGCN: Network in network, dilate convolution and graph convolutional networks based transportation mode recognition[J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(3): 2138–2152. doi: [10.1109/TVT.2021.3060761](https://doi.org/10.1109/TVT.2021.3060761).
- 李宝奇: 男, 1985年生, 特别研究助理, 研究方向为水声信号处理、目标检测、识别和跟踪以及深度学习理论等。
- 黄海宁: 男, 1969年生, 研究员, 研究方向为水声信号与信息处理、目标探测、水声通信与网络等。
- 刘纪元: 男, 1963年生, 研究员, 研究方向为水声信号处理、数字信号处理、水声成像与图像处理等。
- 刘正君: 女, 1982年生, 助理研究员, 研究方向为水声信号处理等。
- 韦琳哲: 男, 1991年生, 助理研究员, 研究方向为水声信号处理等。

责任编辑: 陈倩