

基于改进 CycleGAN 的光学图像迁移生成水下小目标合成孔径声纳图像算法研究

李宝奇^{1,2}, 黄海宁^{1,2}, 刘纪元^{1,2}, 李 宇^{1,2}

(1. 中国科学院声学研究所, 北京 100190; 2. 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 CycleGAN (Cycle Generative Adversarial Networks) 在光学图像迁移生成水下小目标合成孔径声纳图像过程中存在质量差和速度慢的问题, 本文提出一种新的特征提取单元 SDK (Selective Dilated Kernel), 并利用 SDK 设计了一个新的生成器网络 SDKNet. 与此同时, 提出了一种新的循环一致损失函数 MS-CCLF (Multiscale Cyclic Consistent Loss Function), MS-CCLF 增加了图像多尺度结构相似性约束. 在自建的光学图像迁移数据集 OPT-SAS 上, 本文 SM-CycleGAN (Selective and Multiscale Cycle Generative Adversarial Networks) 比原始 CycleGAN 的图像迁移质量提升 4.64%, 生成器网络参数降低 4.13MB, 运算时间减少 0.143s. 实验结果表明, SM-CycleGAN 更适合水下小目标光学图像到合成孔径声纳图像的迁移任务.

关键词: 光学图像迁移生成合成孔径声纳图像; 生成对抗网络; 循环生成对抗网络; 可选择空洞核网络; 多尺度结构相似性

中图分类号: TP391

文献标识码: A

文章编号: 0372-2112(2021)09-1746-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20200712

Optical Image-to-Underwater Small Target Synthetic Aperture Sonar Image Translation Algorithm Based on Improved CycleGAN

LI Bao-qi^{1,2}, HUANG Hai-ning^{1,2}, LIU Ji-yuan^{1,2}, LI Yu^{1,2}

(1. Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China;

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

Abstract: The original CycleGAN show poor quality and time consuming in optical image to underwater small target synthetic aperture sonar image translation task. To address those problems, a novel convolution building block, SDK (Selective Dilated Kernel), is proposed. By stacking SDK blocks, a generator SDKNet is created. At the same time, Multiscale Cycle Consistent Loss Function (MS-CCLF) is proposed, which add the Multiscale Structural Similarity Index (MS-SSIM) between input images and reconstructed images. On our image translation dataset (OPT-SAS), the classification accuracy of our SM-CycleGAN is 4.64% higher than that of original CycleGAN. The generator parameters of SM-CycleGAN is 4.13MB lower than that of CycleGAN, and the time consuming of SM-CycleGAN is 0.143s less than that of CycleGAN. The experimental results show that SM-CycleGAN is more suitable for the translation task of optical image to small underwater target synthetic aperture sonar image.

Key words: optical image-to-synthetic aperture sonar image translation; generative adversarial networks; cycle generative adversarial networks; selective dilated kernel networks; multiscale structural similarity index

1 引言

合成孔径声纳 (Synthetic Aperture Sonar, SAS) 是一种新型高分辨率水下成像声纳^[1,2], 其基本原理是利用

小孔径基阵的移动形成虚拟大孔径, 从而获得方位向的高分辨率. 与普通侧扫声纳相比, SAS 最为显著的优点是方位向分辨率较高, 且理论分辨率与目标距离以

及采用的声波频段无关。不过,获取一幅水下小目标合成孔径声纳图像成本是非常高的^[3,4]。

图像迁移^[5]是计算机视觉一个崭新的研究方向,其目的是实现输入图像与输出图像之间的映射。早期的图像迁移方法尝试利用非参数纹理特征进行人工建模,让迁移的图像能够满足所建立的统计模型。然而这类非参数方法效率很低,并且统计模型的建立也非常困难^[6]。深度学习^[7,8]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^[9,10]。2014年,Goodfellow等^[11]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,不过,GAN在数据生成过程中存在训练不稳定和模式崩溃(Collapse Mode)等问题。为了提高GAN模型的数据生成质量和稳定性,Radford等^[12]提出了深度卷积生成对抗网络DCGAN,它对GAN的体系结构进行了一些修改:将空间池化层函数替换为跨卷积;去除了完全连接层,能提高模型的稳定性;除了生成器的输出层和判别器的输入层之外,对每个单元的输入进行批归一化操作;在生成器中使用ReLU激活函数,在其输出层使用Tanh函数;在判别器中使用LeakyReLU激活函数。为了解决模式坍塌问题,Arjovsky等^[13]提出Wasserstein GAN(WGAN)模型。WGAN使用EM距离代替JS散度来计算生成样本分布与真实样本分布之间的距离。WGAN在生成器和判别器的训练中不需要保持平衡,解决了GAN训练不稳定的问题,同时也解决了模式崩溃现象,保证了生成样本的多样性,并且不需要精细设计网络架构就能实现上述功能。随着GAN理论体系的完善^[14,15],其研究成果却极大地促进图像迁移的发展。2017年,Isola等^[16]人提出了Pix2Pix模型。Pix2Pix首次实现了基于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^[17]人基于GAN和对偶学习设计了循环一致性对抗网络(Cycle Generative Adversarial Networks, 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间的迁移任务。

文献[17]为光学图像迁移生成水下小目标合成孔径声纳图像提供了可能。不过原始CycleGAN生成器网络特征提取层采用固定大小的卷积核(3*3)容易造成

水下小目标信息的缺失。同时,原始循环一致损失函数仅是对输入图像和重构图像的L1约束,换句话说,仅是关于图像颜色的约束,会造成生成图像与真实图像之间存在结构上的差异。对生成器网络而言,通常的做法是增加模型的深度^[18,19]或者增加模型的尺度^[20]。随着对卷积神经网络研究的深入,Li等^[21]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。不过,为了获取多尺度特征,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。对水下小目标图像迁移任务而言,图像的结构信息同样重要。2019年,Huang等^[22]人提出了基于SSIM(Structural Similarity Index)的循环生成对抗网络SS-CycleGAN。SS-CycleGAN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-CycleGAN仅在一个尺度上对输入图像和重构图像进行约束,对水下小目标而言,SSIM无法更准确的表示水下小目标图像在不同尺度上的结构特征。

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低计算量的特征提取单元SDK(Selective Dilated Kernel),并利用SDK模块设计了一个新的生成器网络SDKNet。接着,提出了一种新的基于多尺度结构相似性的循环一致损失函数(Multiscale Cyclic Consistent Loss Function, MS-CCLF)。最后,在CycleGAN框架内通过结合SDKNet和MS-CCLF实现水下小目标光学图像到合成孔径声纳图像的快速、准确的迁移生成。

2 基于多尺度可选择空洞卷积和多尺度结构性约束的水下小目标合成孔径声纳图像迁移生成算法

本文通过改进生成器网络和循环一致损失函数来提高光学图像迁移生成合成孔径声纳图像的质量和速度。

2.1 SM-CycleGAN网络结构

SM-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MS-CCLF组成,网络结构如图1所示。

在SM-CycleGAN上路,光学图像 x 经生成器 G 生成水下小目标合成孔径声纳图像 $G(x)$,再经生成器 F 生成光学重构图像 $F(G(x))$ 。在SM-CycleGAN下路,水下小目标合成孔径声纳图像 y 经生成器 F 生成光学图像 $F(y)$,再经生成器 G 生成水下小目标合成孔径声纳重构图像 $G(F(y))$ 。生成器 $\{G, F\}$ 的网络结构一样,将在2.2节做重点介绍。在SM-CycleGAN中路,判别器 D_x 负责区分

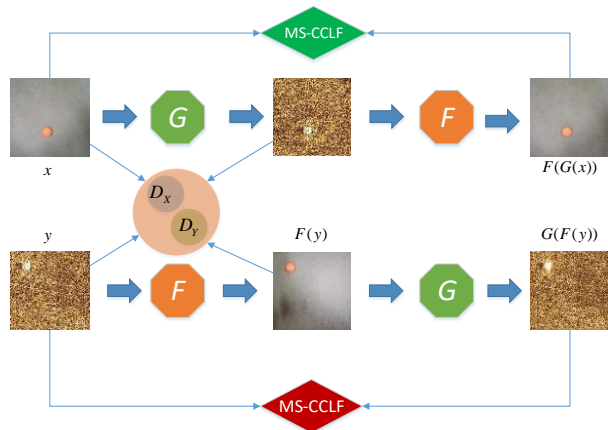


图1 SM-CycleGAN网络结构

光学图像 x 和水下小目标光学生成图像 $F(y)$ 的真假; 判别器 D_y 负责区分水下小目标合成孔径声纳图像 y 和水下小目标合成孔径声纳生成图像 $G(x)$ 的真假. 判别器 $\{D_x, D_y\}$ 的网络结构一样, 均采用 PatchGAN 结构^[17]. 在 SM-CycleGAN 上路和下路的外侧, 循环一致损失函数 MS-CCLF 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 两个域中的数据服从的概率分布. SM-CycleGAN 生成器 $G: X \rightarrow Y$ 的单项损失函数为:

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

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:

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

其中, F 会生成类似于 X 中样本 $F(y)$, D_x 的任务是区分 $F(y)$ 和真实样本 x .

将上面两个损失函数与循环一致损失函数合并可以得到 SM-CycleGAN 总损失函数为:

$$L(G, F, D_x, D_y) = L_{\text{GAN}}(G, D_y, X, Y) + L_{\text{GAN}}(F, D_x, Y, X) + L_{\text{MS-CCLF}}(G, F) \quad (3)$$

其中 $L_{\text{MS-CCLF}}(G, F)$ 表示循环一致损失函数 MS-CCLF, 将在 2.3 节做重点介绍.

SM-CycleGAN 的训练是一个动态的循环过程. SM-CycleGAN 第一阶段, 输入图像 x 通过生成器 G_0 生成图像 $G_0(x)$, 生成图像 $G_0(x)$ 通过生成器 F_0 生成输入图像的重构图像 $F_0(G_0(x))$, 判别器 D_{y0} 对生成图像 $G_0(x)$ 和输入图像 y 进行判断, 结合 $L_{\text{MS-CCLF}}(G, F)$ 调节 $\{G_0, F_0, D_{x0}, D_{y0}\}$ 的参数, 生成第二阶段的 SM-Cycle-

GAN 网络. 第二阶段, 输入图像 y 通过生成器 F_1 生成图像 $F_1(y)$, 生成图像 $F_1(y)$ 通过生成器 G_1 生成输入图像的重构图像 $G_1(F_1(y))$, 判别器 D_{x0} 对生成图像 $F_1(y)$ 和输入图像 x 进行判断, 结合 $L_{\text{MS-CCLF}}(G, F)$ 调节 $\{G_1, F_1, D_{x1}, D_{y1}\}$ 的参数, 生成下一阶段的 SM-CycleGAN 网络. 重复上述过程, 直到得到满意的结果.

2.2 SDK 模块和 SDKNet 生成器网络

SDK 特征提取单元、SK 特征提取单元、bottleneck building block 和 building block 的关系如图 2 所示. 图 2(a) 为 CycleGAN 采用的 ResNet 标准的 building block. 图 2(b) 为 ResNet 改进的 bottleneck building block, 它在图 2(a) 基础上将两个 3×3 的卷积层调整为两个 1×1 的卷积层和一个 3×3 的卷积层, 通过缩放两个 1×1 卷积核的个数降低模型的参数和计算量. 图 2(c) 为 SK 特征提取单元, 它将单一尺度卷积核 (3×3) 替换为多尺度卷积层 ($3 \times 3, 5 \times 5$ 和 7×7), 并利用分组卷积 (groups=32) 提高计算效率, 最后通过 Select 层对多尺度特征进行筛选来提高提取特征的准确性. 图 2(d) 为 SDK 特征提取单元, 它在图 2(b) 和图 2(c) 结构的基础上将大尺寸卷积核 (5×5 和 7×7) 替换为深度可分离空洞卷积^[23], 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^[24]和点卷积, 然后在深度可分离卷积中引入一个称作空洞率^[25, 26]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.

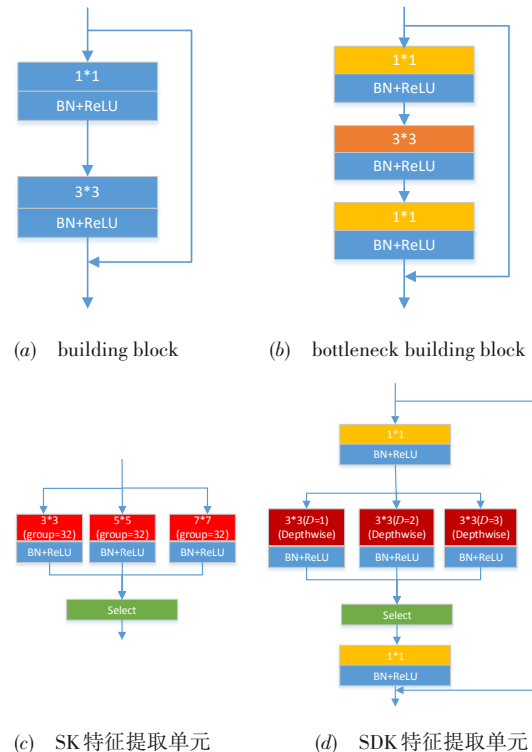


图2 SDK特征提取单元

SDKNet 依然沿用 CycleGAN 原始生成器的网络架构, SDKNet 由输出保持层(第 1 卷积层)、下采样层(第 2~5 卷积层)、特征提取层(第 6~14 卷积层)、上采样层(第 15~18 卷积层)和输出保持层(第 19 卷积层)组成,其中特征提取层(第 6~14 卷积层)用 SDK 模块替换基本特征提取单元模块(building block). 本文 SDK 卷积核尺寸为 3×3 , 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 3×3 、 5×5 、 7×7 和 9×9 四个尺度的卷积核.

以 SM-CycleGAN 上路生成器 G 说明光学图像迁移生成水下小目标合成孔径声纳图像的特征尺寸变化过程. 光学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, 通道数逐层放大一倍、特征尺寸减少一半, 总共四个下采样层, 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 SDK 堆叠组成的特征提取层, 每一层的尺寸和通道数保持不变. 随后, 经过上采样层, 通道数逐层减少一半、尺寸增大一半, 总共四个上采样层, 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小目标合成孔径声纳图像. 生成过程中的生成器 G 和 F 以及图像重构过程中的生成器 G 和 F 中的特征变化过程保持一致.

2.3 基于多尺度结构相似性的循环一致性损失函数

MS-CCLF 新增的多尺度结构相似性函数 MS-SSIM 是一种衡量两幅图像相似度的指标^[27], 它在五个尺度上分别从亮度、对比度、结构三个方面度量图像相似性, 其中均值作为亮度的估计, 标准差作为对比度的估计, 协方差作为结构相似性的度量. 结构相似性的取值范围 $[0, 1]$, 值越大, 表示图像失真越小, 越相似. MS-CCLF 多尺度循环一致损失函数 $L_{MS-CCLF}(G, F)$ 定义为:

$$L_{MS-CCLF}(G, F) = \lambda_{r1} L_{L1}(G, F) + \lambda_{r2} L_{MS-SSIM}(G, F) \quad (4)$$

$$L_{L1}(G, F) = E_{x \sim p_{data}(x)} [\|F(G(x)) - x\|_1] + E_{y \sim p_{data}(y)} [\|G(F(y)) - y\|_1] \quad (5)$$

$$L_{MS-SSIM}(G, F) = E_{x \sim p_{data}(x)} [1 - MS-SSIM(F(G(x)), x)] + E_{y \sim p_{data}(y)} [1 - MS-SSIM(G(F(y)), y)] \quad (6)$$

其中 $L_{L1}(G, F)$ 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; $L_{MS-SSIM}(G, F)$ 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性特征. λ_{r1} 调节原循环一致损失函数, λ_{r2} 调节新增 $L_{MS-SSIM}$ 函数. 事实上, 对于不同的图像迁移生成任务, 图像的迁移生成难度是不同的, 所以选择一个确定的 λ_{r1} 和 λ_{r2} 是非常困难的. 不过考虑 λ_{r1} 和 λ_{r2} 的不同作用, 对于图像背景和结构复杂的图像迁移生成任务,

可以增大 λ_{r2} 的取值; 对于图像背景和结构简单的图像迁移生成任务, 可以减小 λ_{r2} 的取值. 本文 $\lambda_{r1} = 10$ 参考了原 CycleGAN 损失函数的权重, $\lambda_{r2} = 20$ 是通过超参数调优确定.

3 仿真试验

为了验证生成器网络 SDKNet 和循环一致损失函数 MS-CCLF 对 CycleGAN 性能的影响, 实验以水下小目标合成孔径声纳生成图像在 ResNet-18 上分类准确率^[28]、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 CycleGAN 模型为参考, 比较分析 SM-CycleGAN 与其它模型之间的性能差异. 设计实验 2, 比较分析不同生成器网络和循环一致损失函数对 SM-CycleGAN 性能的影响. 设计实验 3, 比较分析不同权重条件下的 MS-CCLF 对 SM-CycleGAN 的影响. 输入图像的尺寸为 284 像素 \times 284 像素, 剪裁后尺寸为 256 像素 \times 256 像素, 特征提取层的个数为 9, 所有模型的 Batchsize 设置为 1.

3.1 水下小目标光学图像-合成孔径声纳图像迁移数据集

为了更好地检验 SM-CycleGAN 对水下小目标合成孔径声纳图像的迁移效果, 我们建立了一个水下小目标光学图像-合成孔径声纳图像迁移数据集: OPT-SAS. OPT-SAS 包括两种水下小目标: 圆形目标和长柱形目标. 水下小目标合成孔径声纳图像与光学图像的具体构成如表 1 所示.

表 1 OPT-SAS 数据集组成

	光学图像	合成孔径声纳图像
圆形目标	76	8
圆柱形目标	72	8
合计	148	16

3.2 实验 1: 图像迁移算法性能比较

本实验比较分析 CycleGAN、SS-CycleGAN、SK-CycleGAN 和 SM-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, 生成器网络基于 building block 模块, 循环一致损失函数采用文献[17]中原始循环一致损失函数, 权重系数等于 10; SS-CycleGAN 表示基于 SSIM 的循环生成对抗网络, 生成器网络基于 building block 模块, 循环一致损失函数采用文献[22]中改进循环一致损失函数, SSIM 项系数等于 0.7; SK-CycleGAN 表示基于 SK 模型的循环生成对抗网络, 生成器网络基于 SK 模块, 循环一致损失函数采用文献[17]中原始循环一致损失函数, 权重系数等于 10; SM-CycleGAN 表示本文提出的循环一致生成器网络, 生成器网络基于本文提出的 SDK 模块, 循环一

致损失函数如式(4)所示,MS-SSIM项权重系数 λ_{r_2} 等于20.模型的迭代次数为400次,前200次采用固定学习率,后200次学习率按线性逐渐衰减至0.保存CycleGAN、SS-CycleGAN、SK-CycleGAN和SM-CycleGAN迭代200次时的生成模型,两类水下小目标光学图像分别经过生成器生成合成孔径声纳图像,使用ResNet-18对两类合成孔径声纳生成图像进行分类识别,并记录分类准确率,参数大小以及运算时间.实验结果如表2所示.

表2 图像迁移算法性能比较

	分类准确率 (%)	参数大小 (MB)	运算时间 (s)
CycleGAN	89.83	10.97	0.154
SS-CycleGAN	89.63	10.97	0.154
SK-CycleGAN	92.54	18.21	0.026
SM-CycleGAN	94.47	6.84	0.011

从表2可以发现,SM-CycleGAN的分类准确率明显高于CycleGAN、SS-CycleGAN和SK-CycleGAN,分别高4.64%、4.84%和1.93%;参数比SK-CycleGAN减少11.36MB,比CycleGAN和SS-CycleGAN减少4.13MB;运算时间比SK-CycleGAN降低0.015s,比CycleGAN和SS-CycleGAN降低0.143s.在参数和运算时间方面,由于SK单元模块的4个多尺度卷积核之间采用的是相加后融合的计算方式,其中每个尺度卷积核的数量与bottleneck building block单尺度卷积核的数量相同,因此参数比原始CycleGAN有明显的增加;同时又由于SK模块采用分组卷积的计算方式,因此运算时间比原始CycleGAN有极大的降低.SM-CycleGAN的SDK模块利用的深度可分离空洞卷积,在降低模型参数的同时也降低了运算时间.综合考虑分类准确率、参数大小和运算时间三个评价指标,SM-CycleGAN优于其他模型.

为了更直观的说明本文算法的性能,将水下圆形小目标光学图像(图3(a))和水下圆形小目标合成孔径声纳图像(图3(d))分别送入SM-CycleGAN的生成模型G和F中,迁移生成声纳图像(图3(b))和生成光学图像(图3(e)),接着利用生成器F和G分别重构光学图像(图3(c))和声纳图像(图3(f)).由水下小目标光学图像(图3(a))迁移生成的声纳图像(图3(b))与真实合成孔径声纳图像(图3(d))的目标特性以及背景纹理非常接近,以达到以假乱真的水平.同时,由真实水下小目标合成孔径声纳图像(图3(d))迁移生成的光学图像(图3(e))与真实光学图像(图3(a))的目标特性以及背景纹理同样非常接近,也达到了假乱真的水平.由此可见,SM-CycleGAN对光学图像迁移生成合成孔径声纳图像的有效性.

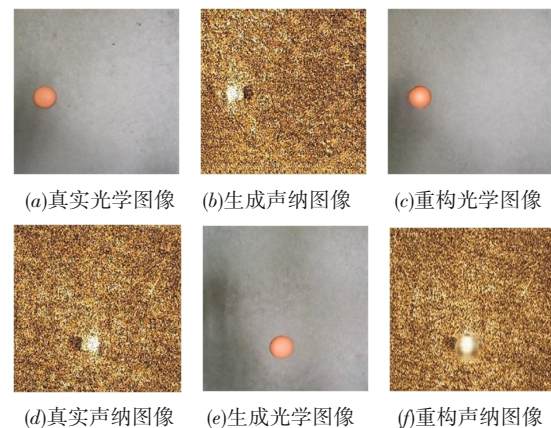


图3 水下小目标合成孔径声纳生成效果

3.3 实验2:生成器网络和循环一致损失函数对CycleGAN的影响

本实验比较生成器网络和循环一致损失函数对CycleGAN的影响.生成器网络分别为RENet、SKNet和SDKNet三种,其中RENet表示生成器网络特征提取层由9个标准building block堆叠组成、SKNet表示生成器网络特征提取层由9个SK模块堆叠组成、SDKNet表示生成器网络特征提取层由9个SDK模块堆叠组成.循环一致损失函数分别为None、SSIM^[22]和MS-CCLF三种,其中None表示不采用结构项约束,SSIM项的权重系数为0.7和MS-CCLF项的权重系数等于20.通过实验组合生成器网络和循环一致损失函数得到不同条件下的CycleGAN模型,模型的迭代次数为400次,前200次采用固定学习率,后200次学习率按线性逐渐衰减至0.保存九种模型迭代400次时的生成模型,两类水下小目标光学图像分别经过上述生成器生成合成孔径声纳图像,使用ResNet-18对两类合成孔径声纳生成图像进行分类识别,并记录分类准确率,实验结果如表3所示.

从表3可以发现,在不采用结构相似性约束条件

表3 生成器和循环一致损失函数对CycleGAN的影响性能的影响

	生成器网络	结构相似性	分类准确率 (%)
CycleGAN	RENet	None	89.83
CycleGAN 变体	SKNet	None	92.54
	SDKNet	None	90.64
	RENet	SSIM	89.63
	SKNet	SSIM	91.46
	SDKNet	SSIM	90.65
	RENet	MS-CCLF	91.94
	SKNet	MS-CCLF	94.71
SM-CycleGAN	SDKNet	MS-CCLF	94.47

下,生成器网络 SKNet 和 SDKNet 优于 RENet,分类准确率分别提升 2.71% 和 0.81%。结构相似性约束 SSIM^[22]并没有为 CycleGAN 带来性能上的提升,RENet、SKNet 和 SDKNet 三种条件下 CycleGAN (SSIM) 与 CycleGAN (None) 的分类准确率基本相同。结构相似性约束 MS-CCLF 提高了 CycleGAN 的性能,与 CycleGAN (None) 相比,RENet、SKNet 和 SDKNet 三种条件下的 CycleGAN (MS-CCLF) 的分类准确率分别提升 2.11%、2.17% 和 3.83%。从 CycleGAN 变体的分类准确率可以看出,基于 SKNet 和 MS-CCLF 的 CycleGAN 变体具有最高的分类准确率,但它的参数大小为 18.21MB 和运算时间为 0.026s (表 2) 比 SM-CycleGAN 要多。SSIM 权重系数等于 0.7 并没有为 CycleGAN 带来性能的提升。事实上,合理选取 SSIM 的权重系数同样能提高 CycleGAN 的性能。综上所述,优化生成器网络和循环一致损失函数均能提高 CycleGAN 的性能。

3.4 实验 3:不同权重条件下的 MS-CCLF 对 SM-CycleGAN 的影响

本实验比较不同权重条件下的 MS-CCLF 对 SM-CycleGAN 的影响,其中生成器采用 SDKNet 网络结构,MS-CCLF 的权重系数 λ_{r2} 分别为 0、0.5、5、10、20、50 和 100,形成不同 7 种不同条件下的 SM-CycleGAN 模型,模型的迭代次数为 400 次,前 200 次采用固定学习率,后 200 次学习率按线性逐渐衰减至 0。ResNet-18 对七种模型下两类水下小目标合成孔径声纳生成图像的分类结果如表 4 所示。

表 4 不同权重条件下的 MS-CCLF 对 SM-CycleGAN 性能的影响

权重系数	0	0.5	5	10	20	50	100
分类准确率 (%)	90.64	89.63	90.44	93.49	94.47	89.22	89.83

从表 4 可以看出,SM-CycleGAN 生成的合成孔径声纳图像分类准确率随 MS-CCLF 权重系数 λ_{r2} 增大呈现先增大后减小的变化趋势,当 MS-CCLF 的权重系数 $\lambda_{r2}=20$ 时,SM-CycleGAN 的分类准确率最高,为 94.47%。同时当 $0 < \lambda_{r2} < 5$ 或 $50 < \lambda_{r2} < 100$ 时,SM-CycleGAN 生成的合成孔径声纳图像分类准确率低于 $\lambda_{r2}=0$ 时的分类准确率,也就是说,当 λ_{r2} 取值过大或过小时并不能给 SM-CycleGAN 带来性能上的提升。

3.5 讨论

实验从分类准确率、参数大小和运算时间三个方面比较了本文图像迁移算法 SM-CycleGAN 与经典算法 (CycleGAN) 和最新算法 (SS-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的选取如何影响 CycleGAN 的性能。CycleGAN 为光学图像迁移生成水下小目标合成孔径声纳图像提供了解决思

路。对于 CycleGAN,性能优异的生成器网络和循环一致损失函数都能提升水下小目标光学图像到合成孔径声纳图像的迁移质量,不过循环一致损失函数并不影响生成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 也验证了这一事实。SDK 模块在 SK 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,同时保存了多尺度特征可选择这一特性,因此性能上优于 SK 模块。MS-CCLF 利用 MS-SSIM 在 5 个尺度上对输入图像与重构图像进行结构约束,相对于单尺度约束,MS-CCLF 的性能更优。对于 λ_{r2} 的权重系数的选取,本文通过实验的方式确定了一个较为合理的数值,事实上, λ_{r2} 权重系数在 [10, 20] 范围内,SM-CycleGAN 均能得到一个性能较优的生成模型。SM-CycleGAN 不仅实现了光学图像迁移生成水下小目标合成孔径声纳图像,重要的是,可以通过水下小目标合成孔径声纳图像迁移生成光学图像增加对声纳图像的识别和理解。

4 结论

光学图像迁移生成水下小目标合成孔径声纳图像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 (SDKNet) 和一种新的循环损失函数 (MS-CCLF),新的生成器网络和新的循环一致损失函数有效提升了 CycleGAN 对水下小目标合成孔径声纳图像的生成质量和效率,经理论分析和仿真实验证明了新方法的有效性。

对水下小目标合成孔径声纳图像迁移任务,改进判别器网络同样能提升 CycleGAN 的性能。下一步的研究重点包括:1) 优化 CycleGAN 判别器网络结构提高模型对真假图像的区分能力;2) 研究更多种类的水下小目标图像迁移模型。

参考文献

- [1] Hayes M P, Gough P T. Synthetic aperture sonar: a review of current status[J]. IEEE Journal of Oceanic Engineering, 2009, 34(3): 207 - 224.
- [2] Wang P, Chi C, Zhang Y, et al. Fast imaging algorithm for downward-looking 3D synthetic aperture sonars[J]. IET Radar, Sonar and Navigation, 2020, 14(3): 459 - 467.
- [3] 刘纪元,唐劲松,孙宝申,等. 基于回波信号的一种合成孔径声纳运动补偿方法[J]. 电子学报, 2003, 31(1): 131 - 134.
- LIU J Y, TANG J S, SUN B S, et al. A receiving-data-based motion compensation method of synthetic aperture sonar[J]. Acta Electronica Sinica, 2003, 31(1): 131 - 134. (in Chinese)

- [4] Sun S B, Chen Y C, Qin L H, et al. Inverse synthetic aperture sonar imaging of underwater vehicles utilizing 3-D rotations[J]. *IEEE Journal of Oceanic Engineering*, 2020, 45 (2): 563 – 576.
- [5] Li Y, Tang S, Zhang R, et al. Asymmetric GAN for unpaired image-to-image translation[J]. *IEEE Transactions on Image Processing*, 2019, 28(12): 5881 – 5896.
- [6] Lin J X, Xia Y C, Qin T, et al. Conditional image-to-image translation[A]. *Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition[C]*. Salt Lake City, USA: Institute of Electrical and Electronics Engineering, 2018. 5524 – 5532.
- [7] Hinton G. Where do features come from? [J]. *Cognitive Science*, 2014, 38(6):1078 – 1101.
- [8] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*, 2016, 521(7553): 436 – 444.
- [9] Schmidhuber J. Deep learning in neural networks: an overview[J]. *Neural Networks*, 2015, 61:85 – 117.
- [10] 贺昱曜, 李宝奇. 一种组合型的深度学习模型学习率策略[J]. *自动化学报*, 2016, 42(6): 953 – 958.
He Y Y, Li B Q. A combinatory form learning rate scheduling for deep learning model[J]. *Acta Automatica Sinica*, 2016, 42(6): 953 – 958. (in Chinese)
- [11] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. *Advances in Neural Information Processing Systems*, 2014, 3: 2672 – 2680.
- [12] Radford A, Metz L. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks[EB/OL]. <https://arxiv.org/pdf/1511.06434v1.pdf>, 2020-07-14.
- [13] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN[EB/OL]. <https://arxiv.org/pdf/1701.07875.pdf>, 2020-07-14
- [14] Chen X, Duan Y, Houthoofd R, et al. InfoGAN: interpretable representation learning by information maximizing generative adversarial nets[A]. *Proceedings of the 2016 Advances in Neural Information Processing Systems [C]*. Barcelona, Spain: Neural Information Processing Systems, 2016. 2172 – 2180.
- [15] Xu Q T, Huang G, Yuan Y, et al. An Empirical Study on Evaluation Metrics of Generative Adversarial Networks [EB/OL]. <https://arxiv.org/pdf/1806.07755.pdf>, 2020-07-14.
- [16] Isola P, Zhu J Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks[A]. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition[C]*. Honolulu, HI, USA: Institute of Electrical and Electronics Engineering, 2017. 5967 – 5976.
- [17] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [A]. *Proceedings of the 2017 IEEE International Conference on Computer Vision[C]*. Venice, Italy: IEEE, 2017. 2242 – 2251.
- [18] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[A]. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition[C]*. Las Vegas, NV, USA: IEEE, 2016. 770 – 778.
- [19] Xie S N, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks[A]. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition[C]*. Honolulu, HI, USA: IEEE, 2017. 5987 – 5995.
- [20] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning[A]. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence[C]*. San Francisco California USA: AAAI Press, 2016. 4278 – 4284
- [21] Li X, Wang W H, Hu X L, et al. Selective kernel networks[A]. *Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition[C]*. Long Beach, CA, USA: IEEE, 2019. 510 – 519.
- [22] Huang X J, Wen L W, Ding J S. SAR and optical image registration method based on improved CycleGAN[A]. *Proceedings of the 2019 Asia-Pacific Conference on Synthetic Aperture Radar[C]*. Xiamen, China: IEEE, 2019. 1 – 6.
- [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.
Li B Q, He Y Y, Qiang W, et al. SSD with parallel additional feature extraction network for ground small target detection[J]. *Acta Electronica Sinica*, 2020, 48(1) 84 – 91. (in Chinese)
- [24] Howard A G, Zhu M L, Chen B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [EB/OL]. <https://arxiv.org/pdf/1704.04861.pdf>, 2020-07-14.
- [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.

- [26] Wang P Q, Chen P F, Yuan Y, et al. Understanding convolution for semantic segmentation[A]. Proceedings of the 2018 IEEE Conference on Applications of Computer Vision [C]. Lake Tahoe, NV: IEEE, 2018. 1451 – 1460.
- [27] Wang Z, Li Q. Information content weighting for perceptual image quality assessment[J]. IEEE Transactions on Image Processing, 2011, 20(5): 1185 – 1198.
- [28] Choi Y, Choi M, Kim M, et al. StarGAN: Unified Generative Adversarial Networks for Multi-domain Image-to-image Translation[EB/OL]. <https://arxiv.org/pdf/1711.09020.pdf>, 2020-07-14.

作者简介



李宝奇 男,1985年出生于天津,中国科学院声学研究所特别研究助理.主要从事水下目标探测、识别和跟踪等方面的研究.

E-mail:libaoqi@mail.ioa.ac.cn



刘纪元 男,1963年出生于辽宁,中国科学院声学研究所研究员,博士生导师.中国图形图像学会视觉与传感专业委员会会员,国家“863”计划“基于无人平台的合成孔径声纳系统研究”项目首席专家.主要研究领域包括水声信号处理、高分辨率水下成像技术等.

E-mail:lji@mail.ioa.ac.cn



黄海宁 (通信作者) 男,1969年出生于河北,中国科学院声学研究所研究员,博士生导师,国务院政府津贴专家.目前担任中科院海洋信息技术创新研究院暨声学研究所科技委副主任,中科院先进水下信息技术重点实验室暨水声工程中心主任,中国声学学会理事.主要从事水声信号与信息处理、目标探测,水声通信与网络等方面的研究.

E-mail:hhn@mail.ioa.ac.cn



李宇 男,1977年出生贵州,中国科学院声学研究所研究员.研究领域涉及水声信号处理、水声通信与网络、无人平台声纳技术、主动声纳技术、阵列信号处理等的多个方面.

E-mail:ly@mail.ioa.ac.cn