

基于图像风格转换的水下图像显著性检测算法

郭继昌¹, 汪昱东¹, 刘迪¹, 艾羽丰¹, 贾伟广²

(1. 天津大学电气自动化与信息工程学院, 天津 300072; 2. 国家海洋标准计量中心, 天津 300112)

摘要: 由于水下显著性检测数据集不足, 导致基于深度学习的水下图像显著性检测网络容易出现过拟合的问题, 从而影响显著性检测网络的性能。针对上述问题, 本文引入图像风格转换方法, 提出一种基于CycleGAN的水下显著性检测网络。网络生成器由图像风格转换子网络和显著性检测子网络构成。首先, 通过无监督的级联方式对风格转换子网络进行风格转换训练, 并利用该网络对陆地图像和水下图像进行风格转换, 构建训练和测试图像数据集, 以解决水下显著性检测数据集不足的问题; 然后, 使用陆地及其风格转换后的显著性数据集对显著性检测子网络进行训练, 以增强网络的特征提取能力; 最后对两个图像风格的输出结果进行融合优化, 以提高显著性检测网络性能。实验结果表明, 本文提出的水下显著性检测网络相比于单纯的陆地和水下图像显著性检测网络, 其检测平均绝对误差和 F 值至少分别提高了10.4%和2.4%。

关键词: 水下图像; 显著性检测; 图像风格转换; 深度学习

中图分类号: TP301.6 **文献标志码:** A

Underwater Image Salient Object Detection Algorithm Based on Image Style Transfer

GUO Jichang¹, WANG Yudong¹, LIU Di¹, AI Yufeng¹, JIA Weiguang²

(1. School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China; 2. National Center of Oceanographic Standards and Metrology, Tianjin 300112, China)

Abstract: Due to the insufficient underwater salient object detection datasets, the underwater image salient object detection network based on deep learning is prone to overfitting, which affects the performance of the network. In response to the above problems, this paper introduces an image style conversion method and proposes an underwater salient object detection network based on CycleGAN. The network generator is composed of an image style conversion subnetwork and a salient object detection subnetwork. First, the network trains the domain transform subnetwork through unsupervised cascade method, and uses the network to perform style transform on in-air and underwater images to construct training and testing datasets, so as to solve the insufficient problem of underwater salient object detection. Then, it uses in-air and salient object detection datasets after style transformation to train the salient object detection subnetwork to enhance the feature extraction ability of the network. Finally, the output results of the two image styles are fused and optimized to improve the performance of the saliency detection network. Experimental results show that compared with the land and underwater salient object detection network, the mean average error (MAE) and F -measure are relatively increased at least 10.4% and 2.4%.

基金项目: 国家自然科学基金(61771334)资助项目。

收稿日期: 2020-11-28; **修订日期:** 2020-12-23

respectively.

Key words: underwater images; salient object detection; image style transformation; deep learning

引 言

图像显著性检测(Salient object detection, SOD)是根据人眼视觉注意力机制,对图像中各部分信息的重要程度进行感知,以获得场景中最吸引注意力的区域。显著图能够有效提高图像中可用信息的提取效率,从而提高如图像分类^[1]和语义分割^[2]等计算机视觉任务的精度,因此显著性目标检测有着重要的研究价值。水下图像显著性检测作为其中一个研究方向,也受到了研究人员的广泛关注。

对于水下图像显著性检测,可以将其分为基于传统方法和基于深度学习方法两类。其中基于传统方法的水下图像显著性检测主要通过颜色、亮度及纹理等低级特征或中心点、边缘等中级特征得到显著性区域。Jian等^[3]用基于四元距离的韦伯描述(Quaternionic distance based Weber descriptor, QD-WD)和模式距离-局部对比算法(Pattern distance and local contrast, PL),提出了一种水下图像显著性检测方法;并在文献^[4]中提出了水下显著性检测数据集,用9种传统显著性检测方法对数据集进行测试,结果显示除PL方法,其余方法的精度、召回率和 F 值(F -measure)均低于50%,这表明传统的显著性检测方法直接应用于成像质量较差、情况较为复杂的水下图像时其检测性能会大幅下降。Chen等^[5]提出了一种新颖的水下显著性目标检测方法,该方法通过增加亮度敏感度、非线性深度敏感度和对比度敏感度,同时考虑3D深度特征,从而获得更准确的显著性图。

近年来,深度学习的快速发展使得基于卷积神经网络的图像显著性检测算法^[6]取得了突破性进展,大量优秀的图像显著性算法不断被提出(如DSS^[7]、R³Net^[8]、PiCANet^[9]和PoolNet^[10])。基于深度学习的显著性目标检测算法对卷积神经网络底层特征和高层特征进行处理,实现边界增强和语义增强,优化了显著性检测的性能,相比传统显著性检测方法有更好的检测准确率。对于水下图像显著性检测,由于缺少相关数据集,使得基于深度学习的水下显著性目标检测算法发展缓慢。Islam等^[11]首次提出了一个可以用于深度学习的水下显著性检测数据集,并提出了Deep SESR网络,对水下图像同时进行图像增强、超分辨率和显著性检测处理。2020年,Islam等^[12]又提出了SVAM-Net,该算法继承了多尺度和语义信息的深层特征,可以在水下图像中进行有效的显著性目标检测。

然而,由于当前水下图像显著性检测数据集较少,难以对基于数据驱动的深度卷积神经网络进行大样本训练,从而无法达到较高的检测精度;同时由于水下图像普遍存在蓝绿色偏、低对比度及亮度不均匀等问题,导致很多性能优异的显著性检测方法在检测退化的水下图像显著性目标时效果不尽人意。针对上述问题,本文将CycleGAN^[13]引入显著性检测网络,提出了一种新颖的端到端的水下显著性检测网络(Underwater salient object detection network, UWSODNet)。该网络包括显著性检测和图像风格迁移两个子网络,通过共享特征提取网络的参数增强两个子网络的关联。本文使用无监督的方式对风格迁移网络进行训练,通过对陆地和水下图像风格进行迁移转换提高水下图像的成像质量,并对陆地图像加入噪声降低其成像效果,有效地减少了陆地和水下图像之间的差异,再使用半监督的方式(仅使用成对的陆地显著性数据集DUTS^[14])对水下显著性检测网络进行训练。然后通过级联结构,对图像进行两次风格迁移并输出两个显著性检测结果,对两个预测图进行叠加以提高显著性检测结果精度。本文提出的UWSODNet能够有效解决数据集不足带来的相关问题。

1 循环生成对抗网络

生成对抗网络(Generative adversarial networks, GAN)^[15]是一种深度学习模型,通过生成器和判别器的不断博弈学习生成质量较好的图像。其中生成器用于生成尽可能真实的假样本,而判别器用于辨别假样本和真实样本,两者通过不断地对抗训练对网络参数进行优化,最终达到纳什均衡点,即判别器不能判别出生成器输出的假样本是否真实。循环生成对抗网络 CycleGAN^[13]作为 GAN 模式的一种,是两个对称的 GAN 构成的环形网络,包含 2 个生成器和 2 个判别器,其目标是学习 X 和 Y 两图像域之间的相互映射关系,其原理图如图 1 所示。图中, G 和 F 为常用的 UNet^[16]网络结构, G 表示从 X 域图像生成 Y 域图像的生成器, F 表示从 Y 域图像生成 X 域的生成器, D_X 和 D_Y 分别表示 X 域和 Y 域的判别器。CycleGAN 通过无监督训练后,能够对不同风格的图像进行相互域转换。在训练中共包含 4 个损失函数:2 个 GAN 损失函数用于对抗生成网络的参数优化,2 个循环损失函数对其循环一致性进行约束,最终实现 $G(F(x)) \approx x$ 和 $F(G(y)) \approx y$ 。

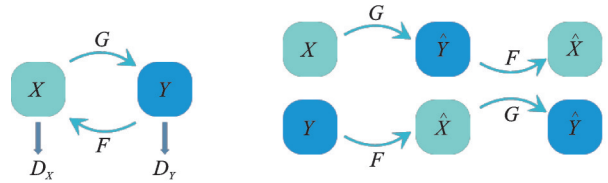


图 1 CycleGAN 原理示意图
Fig.1 Simplified schematic of CycleGAN

2 UWSODNet 网络

2.1 网络结构

本文提出了基于图像风格转换的水下图像显著性检测网络 UWSODNet,通过对陆地图像和水下图像风格进行迁移转换以提高水下图像的成像质量,对陆地图像加入噪声以降低图像的成像效果;通过级联结构,对图像进行两次风格迁移并输出两个显著性检测结果,再对两个预测图进行叠加以提高显著性检测结果精度。UWSODNet 结构示意图如图 2 所示。

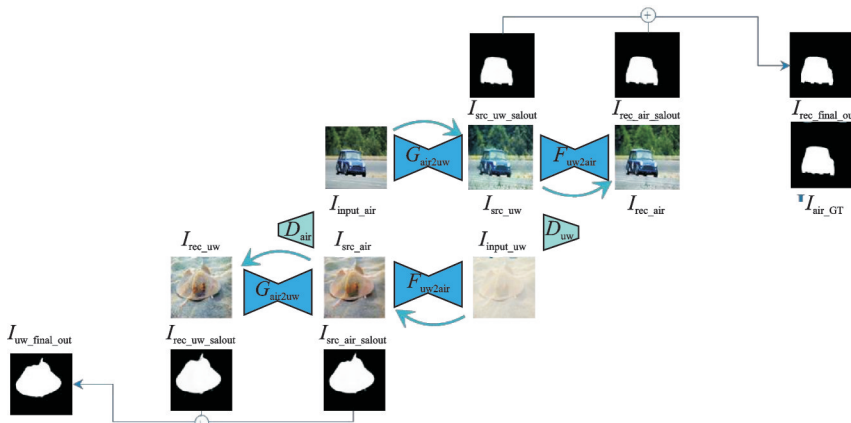


图 2 UWSODNet 网络结构示意图
Fig.2 Network structure of UWSODNet

图 2 中, G_{air2uw} 和 F_{uw2air} 为生成器,用于生成风格转换后的图像和输出显著性检测结果,其中 G_{air2uw} 为陆地图像转换到水下图像的生成器, F_{uw2air} 为水下图像转换到陆地图像的生成器。本文参照文献[17]的方式设计判别, D_{uw} 和 D_{air} 为判别器,分别判断生成的模拟水下图像和模拟陆地图像的真伪。 I_{input_air} 和

I_{input_uw} 分别为陆地和 underwater 真实图像的输入。对于图像风格转换任务, I_{src_uw} 和 I_{src_air} 分别为 I_{input_air} 和 I_{input_uw} 通过 G_{air2uw} 和 F_{uw2air} 生成的模拟水下图像和模拟陆地图像, I_{rec_air} 和 I_{rec_uw} 分别为 I_{src_uw} 和 I_{src_air} 再次通过对抗生成网络获得的还原陆地图像和还原水下图像。对于显著性检测任务, $I_{src_uw_salout}$ 和 $I_{src_air_salout}$ 分别为通过 I_{input_air} 和 I_{input_uw} 获得的显著性检测结果, 而 $I_{rec_air_salout}$ 和 $I_{rec_uw_salout}$ 分别为通过 I_{src_uw} 和 I_{src_air} 获得的显著性检测结果。 $I_{air_final_out}$ 和 $I_{uw_final_out}$ 是分别对上述两个预测图进行叠加而获得的最终陆地和 underwater 图像的显著性检测结果。

UWSODNet 生成器结构是在 CycleGAN 基础上增加了显著性检测子网络, 由于图像风格转换和显著性检测任务均需要强调图像的边缘等细节信息, 因此通过共享特征提取网络的参数增强 2 个子网络之间的关联, 使提取到的特征表现力更强, 从而进一步提升风格转换和显著性检测网络的性能。UWSODNet 生成器网络结构如图 3 所示。

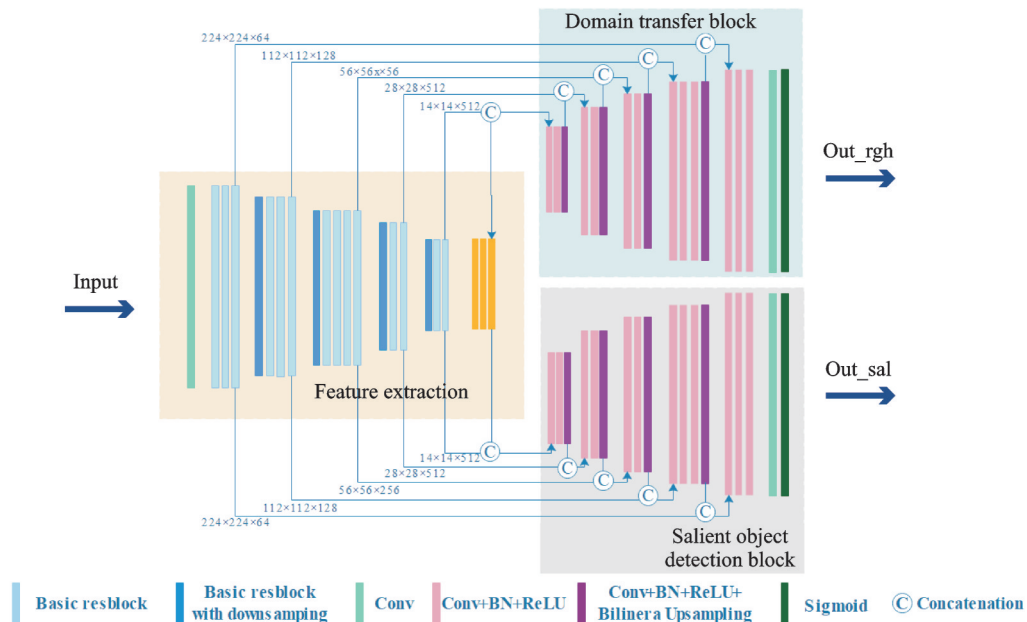


图3 UWSODNet生成器网络结构示意图

Fig.3 Network structure of UWSODNet generator

UWSODNet 生成器包括特征提取、图像风格转换和显著性检测 3 个子网络。其中特征提取网络是下采样网络, 图像风格转换模块和显著性检测网络均为上采样网络。为增加网络特征提取能力, 本文对 UNet 下采样网络进行改进, 使用残差网络结构^[18]替代了原始的卷积操作, 在特征提取最底层增加了 3 个空洞卷积层^[19], 以提高语义信息的提取能力。对于上采样网络, 本文使用 UNet 拼接结构, 用来增加与原图的关联性, 防止编码-解码结构带来的图像信息丢失等问题。特征提取子网络对输入图像进行特征提取, 并同时作为图像风格转换和显著性检测子网络的输入, 这不仅能够提高生成器对图像特征的提取能力, 还能使图像风格转换和显著性检测 2 个子网络的关联性更强, 从而获得更好的风格转换图像和显著性检测图像, 有效提高生成器的性能。本文使用批量归一化 (Batch normalization, BN) 加速网络收敛, 使用 ReLU 作为激励函数, 并使用 Sigmoid 激励函数对输出结果进行处理, 以保证输出图像的像素值保持在 $[0, 1]$ 之间。

2.2 损失函数

对于陆地-水下图像风格转换任务,本文使用无监督的方式对网络进行训练,训练中使用了3种损失函数。包括由水下对抗损失和陆地对抗损失(L_{GANuW} & L_{GANair})组成的 L_{GAN} 、循环损失函数 L_{cyc} 和由水下和陆地结构相似性损失函数(L_{SSIMuW} & L_{SSIMair})组成的 L_{SSIM} ,表达式为

$$\begin{cases} L_{\text{GANuW}}(G, D_{\text{uw}}, X, Y) = E[\log D_{\text{uw}}(y)] + E[\log(1 - D_{\text{uw}}(G(x)))] \\ L_{\text{GANair}}(F, D_{\text{air}}, Y, X) = E[\log D_{\text{air}}(x)] + E[\log(1 - D_{\text{air}}(F(y)))] \\ L_{\text{GAN}} = L_{\text{GANuW}} + L_{\text{GANair}} \end{cases} \quad (1)$$

$$L_{\text{cyc}}(G, F) = E\left[\|F(G(x)) - x\|_1\right] + E\left[\|G(F(y)) - y\|_1\right] \quad (2)$$

$$\begin{cases} L_{\text{SSIMair}}(x) = 1 - \frac{2\mu_x(x)\mu_{G(x)}(x) + c_1}{\mu_x^2(x) + \mu_{G(x)}^2(x) + c_1} \cdot \frac{2\text{Cov}_{xG(x)}(x) + c_2}{\sigma_x^2(x) + \sigma_{G(x)}^2(x) + c_2} \\ L_{\text{SSIMuW}}(y) = 1 - \frac{2\mu_y(y)\mu_{F(y)}(y) + c_1}{\mu_y^2(y) + \mu_{F(y)}^2(y) + c_1} \cdot \frac{2\text{Cov}_{yF(y)}(y) + c_2}{\sigma_y^2(y) + \sigma_{F(y)}^2(y) + c_2} \\ L_{\text{SSIM}} = L_{\text{SSIMair}} + L_{\text{SSIMuW}} \end{cases} \quad (3)$$

式中: x 表示陆地图像; y 表示水下图像; $G(x)$ 表示陆地图像通过生成器 G 生成的模拟水下图像; $F(G(x))$ 表示模拟水下图像通过生成器 F 生成的还原陆地图像;同理 $F(y)$ 和 $G(F(y))$ 分别表示模拟陆地图像和还原水下图像。对于 L_{SSIM} , $\mu(\cdot)$ 、 $\sigma(\cdot)$ 和 $\text{Cov}(\cdot)$ 分别表示以当前像素点为中心、尺寸大小为11像素 \times 11像素的窗口内像素点均值、方差和相对应的协方差。防止分母出现0,本文取 $c_1=0.02$, $c_2=0.03$ 。图像风格转换损失函数 L_{DT} 可以表示为

$$L_{\text{DT}} = L_{\text{GANuW}} + L_{\text{GANair}} + \lambda_1 L_{\text{cyc}} + \lambda_2 (L_{\text{SSIMair}} + L_{\text{SSIMuW}}) \quad (4)$$

对于显著性检测任务,本文仅使用DUTS数据集^[14]对显著性检测网络进行全监督训练,训练中使用了3种损失函数,包括二元交叉熵损失函数 L_{BCE} 、交并比(Intersection over union, IoU)损失函数 L_{IoU} 和结构相似性损失函数 $L_{\text{SSIM}_{\text{SOD}}}$,其表达式为

$$L_{\text{BCE}}(x) = - \sum_x^{H \times W} [S_{\text{gt}}(x) \log(S_{\text{out}}(x)) + (1 - S_{\text{gt}}(x)) \log(1 - S_{\text{out}}(x))] \quad (5)$$

$$L_{\text{IoU}} = 1 - \frac{\sum_x^{H \times W} S_{\text{out}}(x) S_{\text{gt}}(x)}{\sum_x^{H \times W} [S_{\text{out}}(x) + S_{\text{gt}}(x) - S_{\text{out}}(x) S_{\text{gt}}(x)]} \quad (6)$$

$$L_{\text{SSIM}_{\text{SOD}}}(x) = 1 - \frac{2\mu_{\text{gt}}(x)\mu_{\text{out}}(x) + c_1}{\mu_{\text{gt}}^2(x) + \mu_{\text{out}}^2(x) + c_1} \cdot \frac{2\text{Cov}_{\text{gt}(\text{out})}(x) + c_2}{\sigma_{\text{gt}}^2(x) + \sigma_{\text{out}}^2(x) + c_2} \quad (7)$$

式中: $S_{\text{out}}(x)$ 为输出显著图在点 x 的值; $S_{\text{gt}}(x)$ 为真实显著图在点 x 的值。对于 $L_{\text{SSIM}_{\text{SOD}}}$, $\mu_{\text{gt}}(x)$ 、 $\mu_{\text{out}}(x)$ 和 $\sigma_{\text{gt}}(x)$ 、 $\sigma_{\text{out}}(x)$ 分别代表真实显著图和输出显著图在以 x 点为中心、尺寸大小为11像素 \times 11像素的窗口内像素点均值和方差, $\text{Cov}_{\text{gt}(\text{out})}(x)$ 代表其协方差,选取 $c_1=0.02$, $c_2=0.03$ 。由式(5~7)可知 L_{BCE} 为像素级损失函数, $L_{\text{SSIM}_{\text{SOD}}}$ 为区域级损失函数, L_{IoU} 为全局级损失函数。综上,显著性检测损失函数 L_{SOD} 可以表示为

$$L_{\text{SOD}} = L_{\text{BCE}} + L_{\text{SSIM}_{\text{SOD}}} + L_{\text{IoU}} \quad (8)$$

UWSODNet损失函数可以表示为

$$L = \omega_1 L_{\text{SOD}} + \omega_2 L_{\text{DT}} \quad (9)$$

本文设定 $\omega_1=2$ 、 $\omega_2=1$ 、 $\lambda_1=10$ 、 $\lambda_2=5$ 。采用端到端的训练方式,训练时特征提取网络通过 L_{DT} 和

L_{SOD} 进行反向传播更新网络参数,图像风格转换子网络通过 L_{DT} 进行反向传播更新网络参数,显著性检测子网络通过 L_{SOD} 进行反向传播更新网络参数。需要注意的是,由于本文采用了级联的方式,在反向传播时 L_{SOD} 仅更新单个生成器网络参数,而 L_{DT} 不作要求。

3 实验结果分析

3.1 实验方法

本文使用 DUTS 数据集对 UWSODNet 特征提取网络和显著性检测子网络进行全监督训练,选取水下显著性检测数据集 UFO^[11] 和文献[20]提供的水下图像对 UWSODNet 特征提取网络和图像风格转换网络进行无参考训练,并使用自适应矩估计 Adam 作为优化器算法。

训练阶段,本文分 3 步进行训练:首先固定图像风格转换网络的网络参数,对特征提取和显著性检测网络进行预训练,选取学习率(Learning rate, LR)为 0.001,训练回合数 Epoch=10;然后固定特征提取和显著性检测网络的参数,对图像风格模块进行预训练,选取 LR=0.001,Epoch=5;最后在预训练结束后,对 UWSODNet 进行整体训练,选取 LR=0.000 5,Epoch=30。测试阶段,本文仅使用水下图像进行测试。本文所有实验环境均采用 Ubuntu16.04 的设备环境,显卡为 Nvidia GTX 1080Ti(11 GB)。

3.2 评价指标

本文采用两种广泛使用的显著性检测指标来评估实验性能:平均绝对误差(Mean average error, MAE)和 F 值(F -measure)。其中 MAE 表明网络预测的显著图 S_{out} 与真实显著图 S_{gt} 的相似性,表达式为

$$\text{MAE} = \frac{1}{H \times W} \sum_x^{H \times W} |S_{\text{out}}(x) - S_{\text{gt}}(x)| \quad (10)$$

式中 H 和 W 分别表示图像的高度和宽度。MAE 值越小表明预测的显著图效果越好。 F -measure 是一种整体性能度量,通过精度 P 和召回率 R 的加权和平均值计算得出,即

$$F\text{-measure} = \frac{(1 + \beta^2) \times P \times R}{\beta^2 \times P + R} \quad (11)$$

本文选取 $\beta^2=0.3$,使加权精度大于召回率。 F -measure 越大表明网络性能越好。

3.3 实验结果与分析

3.3.1 陆地显著性检测网络对比

本文通过相同的 DUTS 数据集对 UNet^[16]、基于残差网络的 UNet (Res-UNet)、PoolNet^[10]、R³Net^[8]、PiCANet^[9]、DSS^[7] 和本文提出的 UWSODNet 进行训练。由于水下显著性数据集标注效果不佳,本文重新选取了 391 张水下图像进行显著性标注,用于测试上述网络显著性检测效果。实验结果如表 1 所示。

由表 1 第 1、2 列结果可知,使用残差网络结构替换卷积操作可以有效地提高显著性检测的性能。由表 1 第 3~7 列结果可知,当网络通过陆地显著性检测数据集进行有监督训练并预测水下图像时,本文

表 1 使用不同陆地显著性检测网络的水下图像显著性检测结果对比

Table 1 Underwater image detection performance using different in-air salient object detection networks

检测网络	UNet	Res-UNet	PoolNet	R ³ Net	PiCANet	DSS	UWSODNet
MAE	0.136	0.082	0.054	0.057	0.058	0.069	0.048
F -measure	0.640	0.762	0.852	0.845	0.840	0.790	0.879

提出的UWSODNet对水下图像的显著性检测性能最好,其MAE和 F -measure分别为0.048和0.879,相比于陆地图像显著性检测性能较好的网络(PoolNet^[10]、R³Net^[8]、PiCANet^[9]和DSS^[7]),MAE和 F -measure分别有至少0.006(相对12.5%)和0.027(相对3.1%)的提高。使用不同陆地显著性检测网络的水下图像显著性检测部分结果示例如图4所示,其中GT(Ground truth)指图像对应标签。

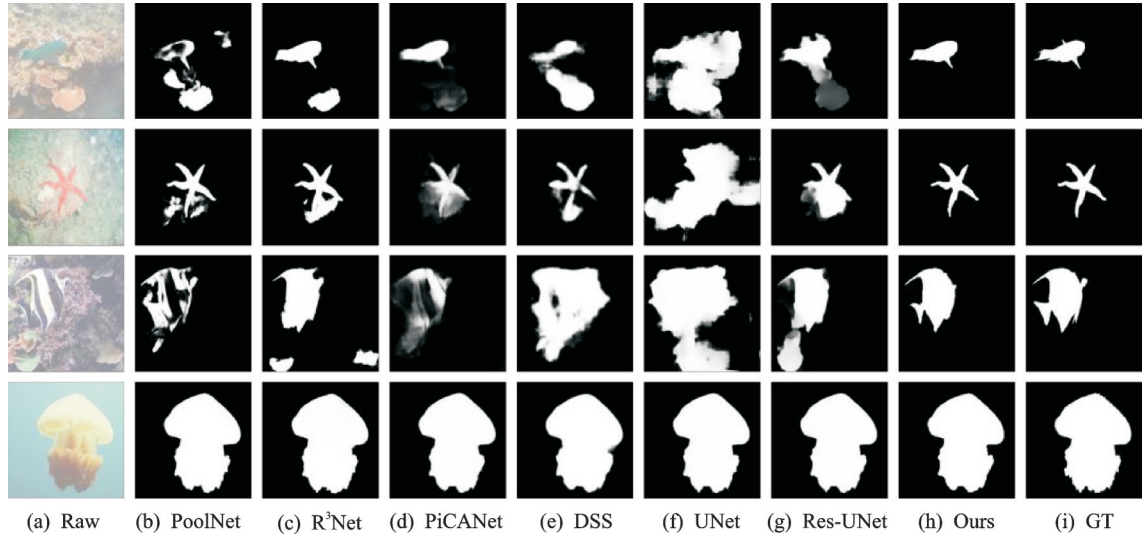


图4 使用不同陆地显著性检测网络的水下图像显著性检测示例

Fig.4 Underwater image detection results using different in-air salient object detection networks

对比图4结果发现,对于背景较为简单的图像,所有算法均能较好地检测到当前水下图像的显著区域。而本文提出的UWSODNet在水下背景较为复杂的图像中能够较好地检测到显著区域,同时其边缘处理得更好。因此可以表明,UWSODNet对水下图像显著性检测的适应性更好,不会过多受到背景的干扰。

3.3.2 水下显著性检测网络对比

为了评估UWSODNet的水下显著性检测性能,本文对比了使用水下显著性数据集进行有监督训练的水下显著性检测网络DeepSESr^[11]和SVAM^[12],测试数据同3.3.1节。本节同时对比了UWSODNet每一个生成器输出的显著图和 $\omega_1=\omega_2=1$ 时网络训练结果,实验结果如表2所示。

表2 使用不同水下显著性检测网络的水下图像显著性检测结果对比

Table 2 Underwater image detection performance using different underwater salient object detection networks

检测网络	DeepSESr	SVAM	$I_{sre_air_salout}$	$I_{rec_uw_salout}$	UWSODNet ($\omega_1=\omega_2=1$)	UWSODNet
MAE	0.372	0.053	0.053	0.074	0.056	0.048
F -measure	0.338	0.858	0.855	0.850	0.845	0.879

由表2可知,在未使用水下显著性检测数据集进行有监督训练的情况下,本文提出的UWSODNet效果优于使用水下显著性检测训练集进行训练的DeepSESr和SVAM,MAE和 F -measure分别提高0.005,0.021(相对为10.4%,2.4%)和0.324,0.541。实验结果表明,本文提出的基于图像风格转换的水

下显著性检测算法能够在没有水下数据集的情况下对水下图像进行有效的显著性检测。由表2第3、4、6列结果可知,本文通过对2个生成器输出的显著图进行叠加能够提高网络检测的性能。对于生成器 F 输出的显著图 $I_{src_aie_out}$,其MAE和SVAM网络相同,说明图像风格转换能够减少陆地图像和水下图像之间的差距,从而提高显著性检测网络的性能。为证明网络损失函数 L 的参数 ω_1 和 ω_2 对检测结果的影响,本文对比了 $\omega_1=\omega_2=1$ 和 $\omega_1=2,\omega_2=1$ 两种训练方式。由表2第5、6列结果可知,在 $\omega_1=\omega_2=1$ 时网络检测效果相较于 $\omega_1=2,\omega_2=1$,MAE和 F -measure分别提高0.008和0.025(相对为16.7%和2.8%)。表2检测结果的部分示意图如图5所示,可见DeepSESR显著性检测效果较差,而本文网络相较于SVAM网络在细节上有一定提高。UWSODNet输出的最终结果相比于2个生成器输出的显著性检测结果 $I_{src_aie_out}$ 和 $I_{rec_uw_out}$ 也有一定的提高。UWSODNet中间输出结果示意图如图6所示。

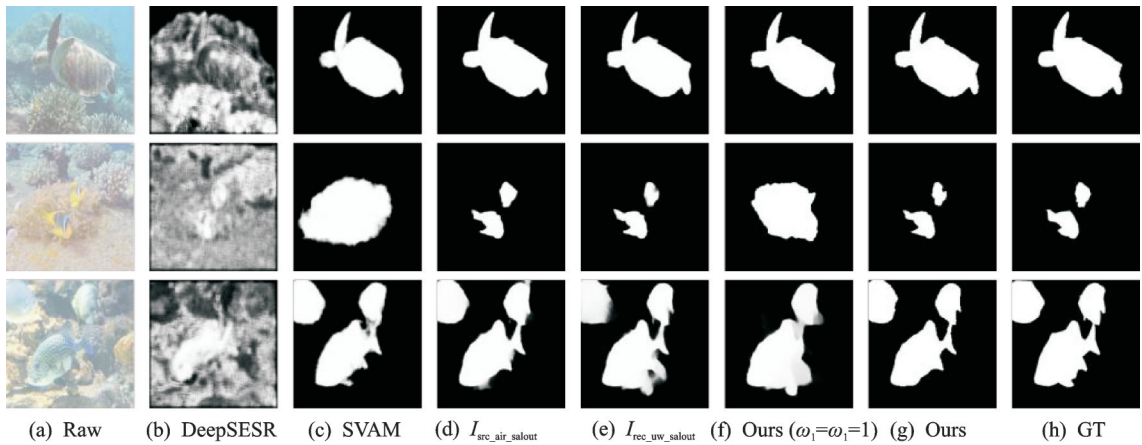


图5 使用不同水下显著性检测网络的水下图像显著性检测示例

Fig.5 Underwater image detection results using different underwater salient object detection networks

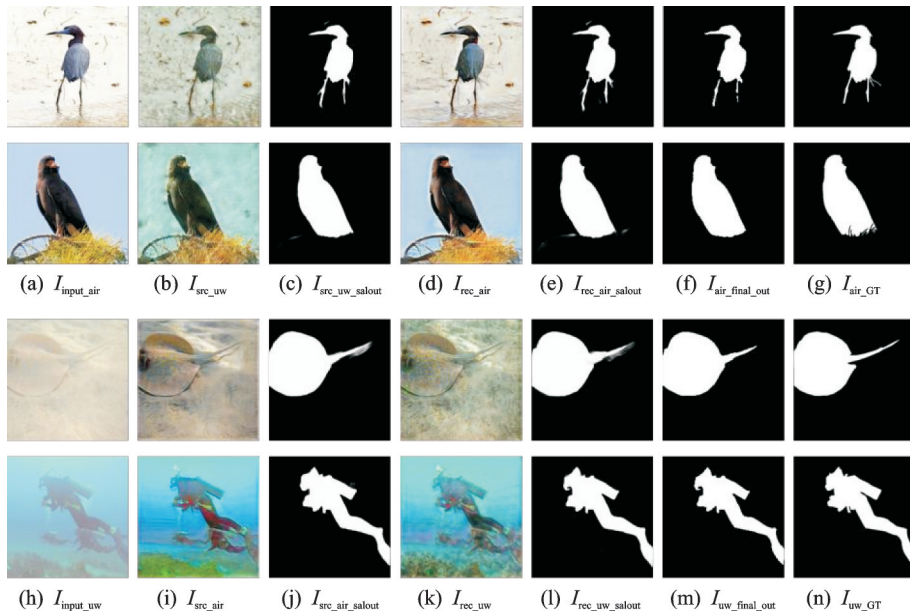


图6 UWSODNet中间输出结果

Fig.6 Intermediate output results of UWSODNet

由图6可知,本文通过图像风格迁移能够使陆地和水下图像风格相互转换,并使陆地图像成像质量下降,水下图像的成像质量提高。通过级联网络输出两次预测结果,并对预测图进行叠加,能够有效提高检测结果的精度。综上,本文提出的UWSODNet在水下图像显著性检测中泛化能力较好,对于背景较为复杂的水下图像也能够较为准确地检测。此外只需要对输入图像进行调整,UWSODNet也可以完成对其他成像质量较差图像(如雾天图像、雨雪天图像)的显著性检测,具有一定的适应性。

4 结束语

本文提出一种新颖的端到端的水下显著性检测网络UWSODNet,将图像风格转换引入了显著性检测网络。该网络能够同时进行图像风格转换和显著性检测两项任务,通过共享特征提取网络的参数增强2个子网络的关联。网络结构基于UNet进行改进,使用残差网络替代了卷积操作并增加了空洞卷积层以提高特征提取能力,同时采用拼接结构用来增加与输入图像的关联性,防止编码-解码结构带来图像信息的丢失。实验结果表明,图像风格转换能够有效降低陆地图像和水下图像之间的差距,能够提高显著性检测网络的性能;通过级联结构,对图像进行两次风格迁移并输出2个显著性检测结果,再对2个预测图进行叠加能够提高显著性检测结果精度。在小样本或零样本的计算机视觉任务中,本文提供了端到端的解决思路,可以对现有数据集进行图像风格转换并对适应当前任务的网络进行训练。然而本文网络的检测准确性和检测速度还有提升空间,下一步将对网络进行改进,做更深入的研究。

参考文献:

- [1] HAN J, NGAN K N, LI M, et al. Unsupervised extraction of visual attention objects in color images[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2005, 16(1): 141-145.
- [2] DING H, JIANG X, SHUAI B, et al. Semantic correlation promoted shape-variant context for segmentation[C]// *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE, 2019: 8877-8886.
- [3] JIAN M, QI Q, DONG J, et al. Integrating QDWD with pattern distinctness and local contrast for underwater saliency detection[J]. *Journal of Visual Communication and Image Representation*, 2018, 53: 31-41.
- [4] JIAN M, QI Q, YU H, et al. The extended marine underwater environment database and baseline evaluations[J]. *Applied Soft Computing*, 2019, 80: 425-437.
- [5] CHEN Z, GAO H, ZHANG Z, et al. Underwater salient object detection by combining 2D and 3D visual features[J]. *Neurocomputing*, 2020, 391: 249-259.
- [6] 梁晔, 马楠, 刘宏哲. 基于深度学习的显著区域提取方法[J]. *数据采集与处理*, 2020, 35(3): 474-482.
LIANG Ye, MA Nan, LIU Hongzhe. Deep learning based salient region detection[J]. *Journal of Data Acquisition and Processing*, 2020, 35(3): 474-482.
- [7] HUO Q, CHENG M, BIRJI A, et al. Deeply supervised salient object detection with short connections[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(4): 815-828.
- [8] DENG Z, HU X, ZHU L, et al. R³net: Recurrent residual refinement network for saliency detection[C]// *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. New Orleans, Louisiana, USA: AAAI Press, 2018: 684-690.
- [9] LIU N, HAN J, YANG M H. Picanet: Learning pixel-wise contextual attention for saliency detection[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, Utah, USA: IEEE, 2018: 3089-3098.
- [10] LIU J J, HOU Q, CHENG M M, et al. A simple pooling-based design for real-time salient object detection[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Long Beach, CA: IEEE, 2019: 3917-3926.
- [11] ISLAM M J, LUO P, SATTAR J. Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception[EB/OL]. (2020-02-04)[2020-11-15]. <https://arxiv.org/abs/2002.01155>.
- [12] ISLAM M J, WANG R B, KARIN L, et al. SVAM: Saliency-guided visual attention modeling by autonomous underwater robots[EB/OL]. [2020-11-15]. <https://arxiv.org/abs/2011.06252>.

- [13] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// Proceedings of IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 2242-2251.
- [14] WANG L, LU H, WANG Y, et al. Learning to detect salient objects with image-level supervision[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, Hawaii, USA: IEEE, 2017: 3796-3805.
- [15] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// Proceedings of Advances in Neural Information Processing Systems. Montréal, Canada: [s.n.], 2014: 2672-2680.
- [16] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. [S.l.]: Springer, 2015: 234-241.
- [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. Honolulu, Hawaii, USA: IEEE, 2017: 1125-1134.
- [18] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 770-778.
- [19] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[C]// Proceedings of International Conference on Learning Representations (ICLR). San Juan, Puerto Rico, USA: [s.n.], 2016.
- [20] FABBRI C, ISLAM M J, SATTAR J. Enhancing underwater imagery using generative adversarial networks[C]// Proceedings of 2018 IEEE International Conference on Robotics and Automation (ICRA). [S.l.]: IEEE, 2018: 7159-7165.

作者简介:



郭继昌(1966-),男,教授,博士生导师,研究方向:智能视频、图像分析、识别及处理,滤波器理论及设计等,E-mail: jcguo@tju.edu.cn。



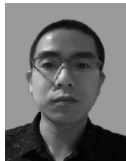
汪昱东(1995-),通信作者,男,博士研究生,研究方向:计算机视觉和图像分析与处理、水下图像增强和复原、图像目标检测、图像显著性检测等,E-mail: yudongwang@tju.edu.cn。



刘迪(1997-),女,硕士研究生,研究方向:图像显著性检测,E-mail: liudi@tju.edu.cn。



艾羽丰(1998-),女,硕士研究生,研究方向:计算机视觉和图像分析与处理,E-mail: aiyufeng@tju.edu.cn。



贾伟广(1981-),男,高级工程师,研究方向:海洋波浪潮汐气象图像处理及计量检测,E-mail: jiaweiguang@ncosm.org.cn。

(编辑:张黄群)