

改进U-Net的水岸分割算法

姚福飞, 柏利志, 周 田

同济大学, 上海

收稿日期: 2022年11月20日; 录用日期: 2022年12月20日; 发布日期: 2022年12月28日

摘 要

我国水域广阔、多样等特点导致难以使用人工的方式实现全面高效地监测、管理, 使用现代信息技术辅助水域管理可以极大地提高效率。水岸分割是水域管理信息化的关键技术, 传统的水岸分割算法依赖手工特征提取需要针对特定的环境调整参数, 适应性不高不适用于多变的天气条件和复杂的河岸场景。基于学习的方法比传统方法更能适应不同的环境, 本文使用基于深度学习的U-Net模型进行水岸分割, 并引入空间注意力模块调整不同空间的特征权重, 使用图像处理方法模拟环境中的雨雾进行数据增强。实验结果表明模型分割MIoU达97.60%, 较已有的算法性能有显著提升。所提模型在极端挑战环境中仍能保持较高的分割准确率, 与传统算法相比具有优秀的环境适应性。

关键词

水域管理, 水岸分割, U-Net, 注意力, 适应性

Improve U-Net for Watershore Segmentation

Fufei Yao, Lizhi Bai, Tian Zhou

Tongji University, Shanghai

Received: Nov. 20th, 2022; accepted: Dec. 20th, 2022; published: Dec. 28th, 2022

Abstract

The characteristics of the vast and diverse waters in the country make it difficult to realize efficient monitoring and management manually. Modern information technology has potential for improving the efficiency of water management greatly. Watershore segmentation is the key technology for water management informatization. Traditional watershore segmentation algorithms rely on manual feature extraction and need to adjust parameters for specific environments, which can't be used in watershore segmentation in a variety of weather conditions and in complex inland river scenarios

due to its poor environment adaptability. Compared with traditional methods, learning-based methods are more adaptable to different environments. In this paper, U-Net is used for the water bank segmentation, which is a deep learning method. The spatial attention module is introduced to adjust the feature weights of different spaces, and the image processing method is used to simulate the rain and fog in the environment for data enhancement. The experimental results show that the model segmentation MIoU is 97.60%, which is significantly improved compared with the existing algorithm. The proposed model maintains high accuracy even in extremely challenging environments, having the property of impressive environmental adaptability compared with the traditional algorithm.

Keywords

Water Management, Watershore Segmentation, U-Net, Attention, Adaptability

Copyright © 2022 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. 背景

随着我国工业化和城镇化的快速发展,水资源与水安全问题日益严重。河流作为城市生态系统的重要组成部分,对社会、人文、生态建设以及景观营造都有着多样的价值[1]。为了加强水环境治理、科学管理河流,需要更准确、快速的监测河流流量以及水质情况。目前河流水量和水质监测主要依靠人工的方式进行,人工的方式受到地理地貌、自然条件、发展状况等因素影响导致误差大、效率低,难以保证准确性和时效性[2]。基于现代人工智能技术的水上无人艇由于具有智能化、无人化、多任务以及实时性等特点,相较于传统人工的方式具有较大优势[3]。国内市场上已有珠海云州智能等多家公司开发可用于水质监测、水下测绘、水文测量和环卫保洁等用途的水上无人艇。相较于在海洋开阔环境,内陆河道具有水文条件复杂、航道狭窄、障碍多等特点,无人智能船在内陆河道平稳运行需要路径规划、障碍躲避以及目标识别等多项技术的支持。水岸分割有助于水上无人艇将水面区域与河岸景观区分开,辅助无人艇进行定位、动态路径规划和障碍躲避,提高水上无人艇的环境适应性[4]。

已有的水岸分割算法基本使用图像处理的方式实现,这些水岸分割算法大致使用特征提取、特征分类与后处理等几个阶段实现水岸分割。基于图像处理的水岸分割算法由于每个处理过程中涉及多个参数,特定环境应用的效果依赖于具体的参数设置,算法的适应性有待改善。近年来快速发展的深度学习技术在计算机视觉领域多项任务中全面超越传统方法,使用深度学习技术的语义分割具有端到端、泛化能力强等特点,相较于传统的基于图像处理的分割方法具有较大优势。

2. 相关工作

2.1. 传统的水岸分割方法

李晓玉等对图像进行多段 Haar 特征提取,根据 Haar 特征大致获取水岸线位置[3]。Fefilyteyev 等使用最大类间差法进行水岸线分割并使用分割后的结果辅助进行水面目标检测[5]。董宇星等使用快速霍夫变换定位海天线然后利用最大熵分割法进行海面小目标检测,该方法适用于形状较规整的海岸环境[6]。郑又能等使用 Lab 颜色空间提取的区域种子点和阈值自适应的生长方法提取水面区域并使用边缘检测的方法提取水

岸分界线[7]。彭明阳等提出采用 HSV (Hue Saturation Value)空间的水面图像特征检测水岸线, 该算法可靠性高, 速度快且效果较好, 但是该算法在水岸线较复杂的区域适用性不高[8]。采用传统图像识别的方法利用水面与地面的一些特征进行图像分割, 针对不同的场景需要手动调整参数设计特征, 适用性不强[4]。

2.2. 深度学习方法

基于深度学习的图像语义分割方法对图像上每一个像素进行分类从而实现区域的识别, 2012 年 Krizhevsky 等人提出的 AlexNet 引入多项技术解决了深度神经网络的梯度消失和过拟合问题, 在 2012 年 ImageNet 比赛中以远超当时亚军的优势获得分类任务冠军[9], 自此深度学习开始被大量应用于各类计算机视觉任务中。FCN (Fully Convolutional Networks)使用卷积和池化得到一系列特征图, 使用反卷积和跳连接得到精细的语义分割结果, 因整个网络不包含任何全连接层而命名全卷积网络[10]。U-Net 通过使用卷积和最大池化进行特征提取, 解码器堆叠反卷积的特征图和从编码器传过来的低级特征图后卷积得到语义分割结果, 通过数据增强手段训练后的网络在训练数据集较少的情况下以较大优势获得 2012 年国际医疗图像分割挑战赛冠军[11]。Deeplab v3+使用深度可分离卷积进行特征提取和融合, 在保持特征高维的同时减少了模型参数量, 融合多尺度空间金字塔池化提取的特征使每一个分类结果均包含全局信息[12]。沈建军等在 Deeplab v3+解码器中多加一路低级特征通路增加了分类的特征信息, 并将其应用于水岸检测任务中[4]。PSPNet 中使用空洞卷积和全局池化技术增大卷积感受野, 并使用金字塔特征池化产生的多尺度特征图作为全局场景先验信息避免了信息丢失, 有效地提高了小物体的分割检测能力[13]。Yin 等在 PSPNet 中引入空洞卷积和特征金字塔池化, 并使用迁移学习方法训练了一个准确率较高的模型[14]。Erdem 等人提出了基于不同神经网络结构的多数投票方法自动获取水岸线[15], 基于深度学习的方法对不同的水岸线场景具有适应性且能够克服诸如反射和水波等不可控因素的干扰[16][17]。

注意力被视作可以根据输入信号改变计算资源分配的机制。Jaderberg 等在需要学习平移、缩放以及旋转不变性的卷积神经网络中使用可微分的空间注意力模块提高了模型性能[18]。Hu 等提出 SE (Squeeze & Extention)模块自适应地对特征图不同位置的关系进行建模和矫正, 使用 SE 模块的 ResNet 在图像分类和目标检测等任务中取得一定的性能提升[19]。

Borja 等开源的 MaSTr1325 大规模海洋语义分割训练数据集, 用于开发小型沿海水面无人艇的障碍物检测方法[20]。Cheng 等发布了国内首个用于内河水面上无人艇的数据集 USVInland, 该数据集包含了用于水面语义分割的图像和对应标签、用于机器人定位的传感器数据以及用于立体匹配的数据集, 是一个专为开发内河水面上无人艇而设计的数据集[21]。

本文借鉴深度学习方法语义分割中的 U-Net 结构[11], 引入并探索[19]所提的 SE 模块在提高水岸分割精度以及改善模型鲁棒性方面的作用。

3. 算法原理与实验

3.1. 模型结构

U-Net 网络结构是典型的编码器 - 解码器型结构。编码器部分使用步长为 2 的最大池化对输入特征图进行下采样, 下采样得到的特征图经过多次卷积后使用 SE 模块对特征图进行矫正, 对输入图像进行多次下采样、卷积和矫正操作将图像编码成高维表示。解码器使用转置卷积对特征图进行上采样, 上采样得到的特征图与编码器通过跳连接(Skip Connection)传递的特征图堆叠进行多次卷积, 对卷积后的特征图使用 SE 模块进行空间矫正, 高维特征图通过多次上采样、卷积和矫正操作逐步恢复原图像大小。对解码器输出的大小与原图相同、通道数为 2 的特征图应用 Softmax 激活函数后分别表示图像上对应像素是水域和陆地的概率, 加入 SE 模块的 U-Net 结构如图 1 所示。

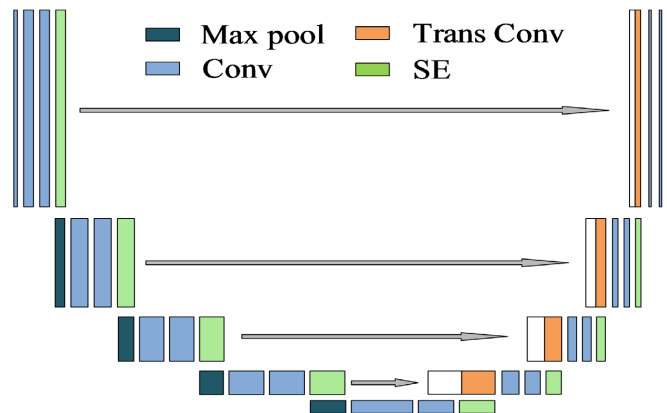


Figure 1. U-Net architecture with SE module
图 1. 加入 SE 模块的 U-Net 结构

SE 模块对输入特征图进行通道压缩和空间激活操作，图 2 是 SE 模块的结构示意图。假设输入特征图为 $U \in \mathbb{R}^{C \times H \times W}$ ，首先对特征图应用卷积核大小为 1×1 的卷积将输入特征图压缩成 $U' \in \mathbb{R}^{1 \times H \times W}$ ， U' 中每一个数字 u_{ij} 均为特征图在空间位置 (i, j) 处 C 个通道的值的线性组合， U' 经过非线性激活函数 $\sigma(\cdot)$ 调整到 $[0, 1]$ 区间后再同原输入特征图 $U \in \mathbb{R}^{C \times H \times W}$ 相乘进行空间矫正。激活后的每一个值 $\sigma(u_{ij})$ 表示给定特征图 (i, j) 处的空间信息的相对重要性，空间矫正操作将相关的空间位置的重要性提高并且降低不相关位置的特征图的重要性。

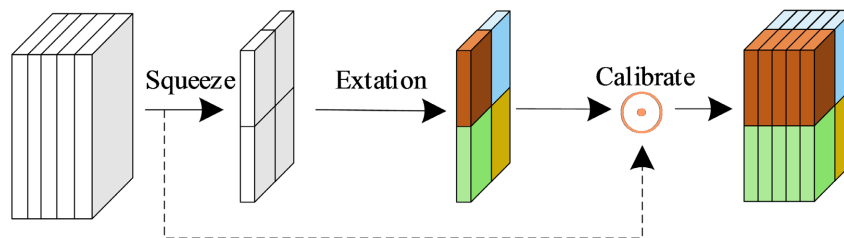


Figure 2. SE module
图 2. SE 模块

3.2. 损失构成

训练损失由逐像素的交叉熵损失 L_{CE} 和度量预测水面区域与真实水面区域交叠情况的戴斯损失 L_{Dice} 构成，戴斯损失(Dice Loss)是使用两个集合的交集的元素个数与两个集合元素个数的总和的比值度量两个集合的相似性。最终损失函数的数学形式为：

$$L = L_{CE} + \alpha L_{Dice}, \text{ 其中 } L_{CE} = - \sum_{y \in Y, \hat{y} \in \hat{Y}} y \log(\hat{y}), L_{Dice} = 1 - \frac{2|Y \cap \hat{Y}|}{|Y| + |\hat{Y}|}$$

其中 Y 表示真实水面区域的像素集合， \hat{Y} 表示模型预测的水面区域的像素集合， α 为调节两项损失的权重因子。

3.3. 图像增强

由于实际作业环境中存在光照、雨雾以及季节变化等多种因素影响水上无人艇运行，本文使用多种数据增强手段进行数据预处理，以期提高模型在这类环境中的鲁棒性。具体使用的数据增强手段包括模

拟光照变化的亮度饱和度变换、模拟雨雾的天气变换、模拟不同水面区域界线的仿射变换以及裁剪平移等变换¹。将图像转换成 HSV 颜色空间后调整相应通道实现图像亮度和饱和度变换，通过在图像上用白色画出一系列方向相同的线条模拟传感器观测到降雨的图像，将图像的三通道与随机生成的白色图像进行线性混合可以实现模拟雾观测的效果。图 3 是对某个训练样例应用本实验中使用的最强的亮度变换、模拟降雨和模拟雾环境的效果图，左上、右上、左下和右下分别是原图、亮度调整后的图像(明度通道乘以 0.8)、模拟降雨(Density = 0.18)的图像和模拟起雾(Intensity = 0.7)的图像。



Figure 3. Data argument sample
图 3. 数据增强样例

4. 结果与分析

4.1. 模型分割性能

本文使用 USVInland 的水岸分割数据集进行模型训练，该数据集一共含有 466 个用于训练的语义分割图像标注，234 个用于测试的语义分割图像标注。训练使用图像尺寸为 400×640 ，每一个批量输入 8 个训练样例。初始学习率为 0.001，学习率策略为发现 L_{Dice} 连续多次不下降时将学习率除以 10。图 4 为本文所提 SE-U-Net 模型和原 U-Net 模型前 30 次迭代后在测试集上的平均交并比(Mean Intersection over Union, MIoU)变化情况，通过 SE 模块对特征图进行加权矫正利于模型选择用于分类的有用信息，所以从初始训练开始含有 SE-U-Net 模型测试 MIoU 上升较快。图中拐点为 MIoU 不再提高时调整学习率形成，无论是学习率调整之前还是学习率调整之后，SE-U-Net 模型的最大 MIoU 比原始 U-Net 模型均有提高，这佐证了引入 SE 模块之后的模型比 U-Net 拥有更好的分类能力。

表 1 是参考文献[12][13][14]的所述的模型与本文所提模型在 USVInland 数据集上的分割性能指标，其中背景项代表分割结果计算的除水面之外的其他区域的 MIoU，前景项代表分割结果计算的水面区域的 MIoU，平均项是二者的平均值。SE-U-Net 模型性能比 DeepLab v3+、PSPNet 和原始 U-Net 模型均有不同程度的提升，仅[14]使用多项技术改进的 PSPNet 模型在背景识别上稍好于本文所提模型，但是本文的模型在前景水面区域的识别方面有较大优势，最终 SE-U-Net 水面区域识别的 MIoU 和平均 MIoU 均优于[14]所提模型。

¹ 图像增强相关操作请参考 <https://github.com/aleju/imgaug>。

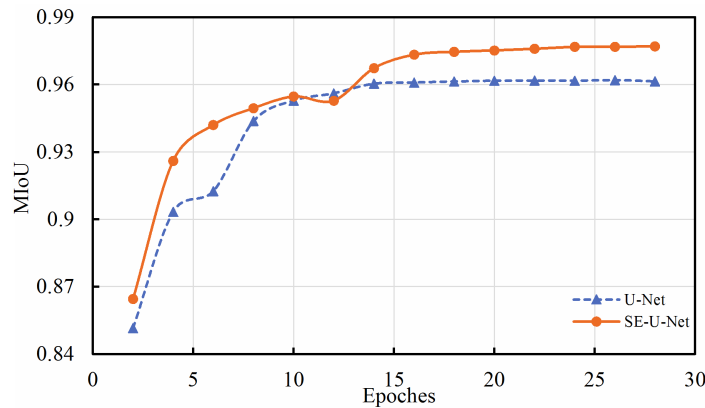


Figure 4. MIoU during training

图 4. 训练 MIoU 曲线图

Table 1. Background and foreground MIoU of different models

表 1. 不同模型背景和前景 MIoU

模型	背景	前景	平均
DeepLab v3+	96.09%	93.47%	94.01%
PSPNet	96.87%	95.02%	95.94%
Improve PSPNet	97.59%	96.14%	96.87%
U-Net	95.12%	96.02%	95.69%
SE-U-Net	97.10%	97.60%	97.41%

4.2. 挑战环境下的模型性能

为了验证模型的适应性, 本文在测试数据集的基础上模拟了雨和雾等挑战因素来测试模型环境适应性。一定浓度的雨和雾影响下的模型分割 MIoU 结果如图 5 所示, 其中左边是受到降雨影响的模型分割 MIoU 对比, 右图是受到雾影响的模型分割 MIoU 对比图。从图中可看出随着影响因素浓度增大, SE-U-Net 模型和 U-Net 模型分割的 MIoU 均有所下降, 但总体上 SE-U-Net 模型性能下降比 U-Net 下降慢, 环境适应性较 U-Net 模型有提升。在模拟的下雨环境下, 两个模型分割性能受到的影响较小, 但是在模拟雾的环境中两个模型性能差异较大。U-Net 模型在受到雾的影响时模型分割性能迅速下降, 而 SE-U-Net 在雾浓度为 0.7 的测试集上 MIoU 接近 90%, 即 SE-U-Net 模型在受到最浓的雾影响时模型性能尚可勉强接受, 对环境有一定的适应能力。

实际工作环境不仅有上述常见光照、天气变化的影响, 而且有很多诸如相机镜头污垢、倒影以及复杂环境等因素影响。当相机镜头上有污垢时, 图像往往在局部区域呈现一片光晕或斑点; 倒影则由于与岸上景观具有相似的结构、纹理等因素非常容易错判, 部分颜色较深的物体(如树木)形成倒影时图像水岸线两侧区分并不明显, 难以精确地分割水岸区域。这些影响因素常常叠加在同一帧图像上出现, 叠加的多个影响因素对模型的环境适应性提出了非常高的要求。图 6 是本文测试的 SE-U-Net 模型和 U-Net 模型在实际影响因素的作用下的分割识别效果, 图像上的影响因素从上到下分别是正常状态、镜头污垢和降雨、少量污垢和雾、极端亮度变化和倒影, 从左到右分别是原始 RGB 图像、真实类别标签(Ground Truth)、U-Net 模型分割标签、SE-U-Net 模型分割标签。在正常状态下两个模型均能较好地进行图像识别和语义分割; 第二帧图像镜头上的镜头污垢明显对 U-Net 模型的分类和分割造成了较大影响, 图像右侧边缘污

垢形成的斑点被模型误识别成水面区域；第三帧图像中的雾对 U-Net 模型影响较大，图像左侧部分树木区域由于局部区域亮度较低和雾的影响已经难以区分树木边界，U-Net 模型将这一部分区域识别为水面区域造成分割结果与真实水面区域出现较大差异，而 SE-U-Net 模型相对受到的影响较小，识别结果中仅少量边界被误识别；第四帧图像中两者识别的水面区域差距不大，细节上 U-Net 模型倾向于将有倒影的水面区域识别背景区域，而亮度较低的背景区域有可能被 U-Net 识别成水面区域，SE-U-Net 模型受到的影响较小，能更好地识别水面倒影区域和岸上的低亮度区域。

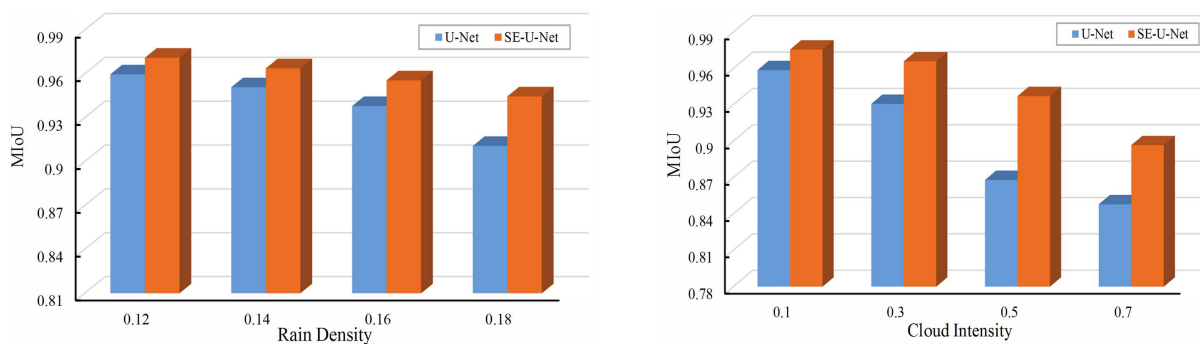


Figure 5. MIoU under the influence of rain and cloud

图 5. 雨雾影响下的模型 MIoU

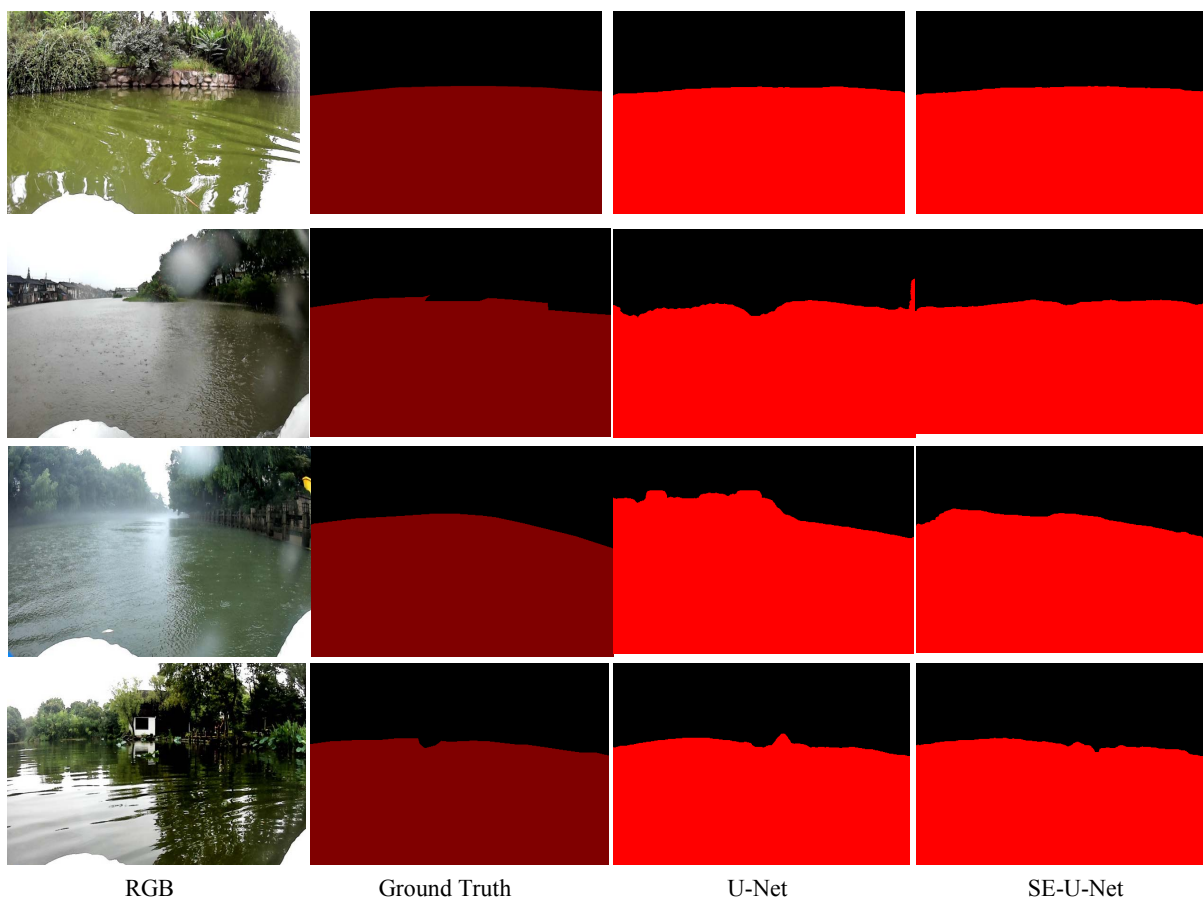


Figure 6. Semantic label of challenging environment

图 6. 挑战环境分割语义标签

5. 结论

本文在 U-Net 模型中加入 SE 模块的方法能够有效提高模型水岸分割性能,并结合相应的数据增强方法进行训练,最终的模型在测试数据集上 MIoU 较已有方法有较大提高,并且能够有效应对雨雾、倒影、传感器污垢以及光照变化等因素对模型分割的影响,在极端的环境下较准确的水面区域识别。该模型在水域监控和水上无人艇视觉导航方面具有一定实用价值。但是该模型计算量庞大,只能在 GPU 设备上运行,还无法适应小型设备上的应用,研究如何降低模型复杂度有助于在移动嵌入式设备中应用水岸分割模型。

参考文献

- [1] 宋庆辉, 杨志峰. 对我国城市河流综合管理的思考[J]. 水科学进展, 2002, 13(3): 377-382.
- [2] 李晓玉, 王鹏, 邵光艺, 等. 基于水面无人船快速监测的温榆河-北运河水质分析与评价[J]. 环境化学, 2022, 41(5): 1568-1578.
- [3] 鲍江, 陶青川, 张鹏. 基于图像处理的水位线检测算法[J]. 水电能源科学, 2015, 33(4): 96-99,210.
- [4] 沈建军, 陶青川, 肖卓. 结合改进 Deeplab v3+网络的水岸线检测算法[J]. 中国图象图形学报, 2019, 24(12): 2174-2182.
- [5] Fefilatyeve, S. and Goldgof, D. (2008) Detection and Tracking of Marine Vehicles in Video. 2008 *IEEE 19th International Conference on Pattern Recognition*, Tampa, 8-11 December 2008, 1-4. <https://doi.org/10.1109/ICPR.2008.4761344>
- [6] 董宇星, 刘伟宁. 基于灰度特性的海天背景小目标检测[J]. 中国光学与应用光学, 2010, 3(3): 252-256.
- [7] 郑又能, 肖宇峰, 寸超, 等. 基于改进区域生长的水岸线提取方法研究[J]. 计算机应用研究, 2020, 37(6): 1876-1881.
- [8] 彭明阳, 王建华, 闻祥鑫, 丛晓奕. 结合 HSV 空间的水面图像特征水岸线检测[J]. 中国图象图形学报, 2018, 23(4): 526-533.
- [9] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. <https://doi.org/10.1145/3065386>
- [10] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 3431-3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- [11] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Berlin, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [12] Chen, L.-C., Zhu, Y., Papandreou, G., et al. (2018) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, 8-14 September 2018, 801-818. https://doi.org/10.1007/978-3-030-01234-2_49
- [13] Zhao, H., Shi, J., Qi, X., et al. (2017) Pyramid Scene Parsing Network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 2881-2890. <https://doi.org/10.1109/CVPR.2017.660>
- [14] Yin, Y., Guo, Y., Deng, L. and Chai, B. (2022) Improved PSPNet-Based Water Shoreline Detection in Complex Inland River Scenarios. *Complex & Intelligent Systems*, 1-13. <https://doi.org/10.1007/s40747-022-00793-8>
- [15] Erdem, F., Bayram, B., Bakirman, T., et al. (2021) An Ensemble Deep Learning Based Shoreline Segmentation Approach (WaterNet) from Landsat 8 OLI Images. *Advances in Space Research*, **67**, 964-974. <https://doi.org/10.1016/j.asr.2020.10.043>
- [16] von Braun, M.-S., Frenzel, P., Kading, C. and Fuchs, M. (2020) Utilizing Mask R-CNN for Waterline Detection in CANOE Sprint Video Analysis. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Seattle, 14-19 June 2020, 876-877. <https://doi.org/10.1109/CVPRW50498.2020.00446>
- [17] McLeay, A.J., McGhie, A., Briscoe, D., et al. (2021) Deep Convolutional Neural Networks with Transfer Learning for Waterline Detection in Mussel Farms. 2021 *IEEE Symposium Series on Computational Intelligence (SSCI)*, Orlando, 5-7 December 2021, 1-8. <https://doi.org/10.1109/SSCI50451.2021.9659987>
- [18] Jaderberg, M., Simonyan, K. and Zisserman, A. (2015) Spatial Transformer Networks. *Proceedings of the 28th International Conference on Neural Information Processing Systems*, Volume 2, 2017-2025.

-
- [19] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>
- [20] Bovcon, B., Muhovič, J., Perš, J. and Kristan, M. (2019) The MaSTr1325 Dataset for Training Deep USV Obstacle Detection Models. 2019 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, 3-8 November 2019, 3431-3438. <https://doi.org/10.1109/IROS40897.2019.8967909>
- [21] Cheng, Y., Jiang, M., Zhu, J. and Liu, Y. (2021) Are We Ready for Unmanned Surface Vehicles in Inland Waterways? The USVInland Multisensor Dataset and Benchmark. *IEEE Robotics and Automation Letters*, **6**, 3964-3970. <https://doi.org/10.1109/LRA.2021.3067271>