

# 面向滨海生态监管的多尺度目标语义分割研究

陈 岩<sup>1</sup> 杨晓彤<sup>2</sup> 奚砚涛<sup>3</sup> 徐立祥<sup>1</sup> 李新路<sup>1</sup>

1. 合肥学院人工智能与大数据学院, 合肥;

2. 国家海洋信息中心海洋测绘地理信息部, 天津;

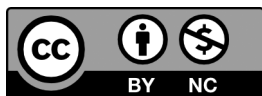
3. 中国矿业大学资源与地球科学学院, 徐州

**摘 要** | 针对缺少滨海生态场景深度学习数据集, 面向遥感影像分类的多尺度目标语义分割精度不高等问题, 研究以红树林、浮筏养殖和围塘养殖三类滨海典型生态监管多尺度目标为研究对象, 构建了面向滨海生态监管的多目标语义分割数据集, 通过集成批归一化和空间置弃算法, 改进 UNet 特征融合策略, 提出了一种多尺度深度卷积语义分割模型。模型在测试集上总体精度 92%, Kappa 系数 0.87, 平均交并比 82%。实验结果表明批归一化与特征融合空间置弃的耦合堆叠可有效抑制多尺度目标语义分割过拟合, 提高模型精度和泛化性能。研究提出的模型及构建的面向滨海生态环境监管的多目标语义分割数据集可为滨海区域生态修复、测绘和综合治理提供决策支持。

**关键词** | 滨海生态监管; 人工智能; 红树林; 海水养殖; 基准数据集; 多尺度特征融合; 语义分割

Copyright © 2022 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). <https://creativecommons.org/licenses/by-nc/4.0/>



基金项目: 安徽省教育厅高校自然科学基金 (KJ2020A0658); 合肥学院科学研究发展基金 (20ZR03ZDA); 合肥学院人才科研基金 (20RC13)。

通讯作者: 陈岩, 安徽六安人, 讲师, 博士, 研究方向: 遥感图像处理与人工智能研究, Email: chenyan090501@126.com。

文章引用: 陈岩, 杨晓彤, 奚砚涛, 等. 面向滨海生态监管的多尺度目标语义分割研究 [J]. 环境与资源, 2022, 4 (2): 48–61.

<https://doi.org/10.35534/er.0402007>

## 1 引言

随着海产品需求的增长和利益驱使,无序、非法海水养殖活动日益频繁,不仅对红树林、泥炭沼泽和珊瑚礁等构成的滨海湿地生态环境造成了破坏,也给滨海区域规划及其综合治理带来了极大挑战<sup>[1]</sup>。遥感以其探测范围广、时效性强又经济等特点,为滨海区域监测提供了强有力支持。借助遥感影像和分类算法自动、快速、准确地提取目标地物是遥感领域的重要课题之一,也是有效监管和保护滨海生态环境的重要手段。

传统滨海区域遥感监测多采用中低空间分辨率光学影像<sup>[2-5]</sup>或雷达影像<sup>[6-8]</sup>。随着传感器和航空航天技术的发展,包含丰富空间信息的高分辨率遥感影像数量增多,成本也在下降,受到国内外学者关注。为从高分辨率遥感影像中准确提取滨海生态监管目标,研究人员尝试了各种技术方法,例如基于光谱的图像分类<sup>[9]</sup>和面向对象的图像分析<sup>[10]</sup>。然而,丰富的空间信息需要复杂的光谱分布响应,而高分辨率遥感影像的低光谱分辨率导致目标地物在光谱域的可分性不高;而面向对象的图像分析方法需要首先基于纹理、形状等特征对图像进行超像素分割。由于缺乏统一标准,容易引入新的误差并传递至分类阶段,影响整体性能。

近年来,以 AlexNet<sup>[11]</sup> 为代表的深度卷积神经网络 (deep convolutional neural network, DCNN) 在计算机视觉领域取得了巨大成功,被看作是新一代人工智能技术爆发的起点。在遥感领域,基于 DCNN 的各种深度学习模型在土地利用覆盖场景分类、船只与飞机检测、建筑与道路提取等应用中取得了良好效果<sup>[12]</sup>。它们分别采用了计算机视觉中的图像识别、目标检测和语义分割方法。而术语“语义分割”与“遥感影像分类”或“提取”的目标一致,实现像素级分类。全卷积网络 (fully convolutional network, FCN)<sup>[13]</sup> 是 DCNN 在图像语义分割应用中的代表性模型。它将经典 DCNN 的卷积层保留作为特征提取器,将全连接层替换为卷积层,并引入插值方法恢复图像空间信息。与面向中低分辨率遥感影像分类常用的基于图块 (Patch-based)<sup>[14]</sup> 的 DCNN 方法相比,FCN 实现了端到端语义分割。此后,以 FCN 为基础设计的 U-Net<sup>[15]</sup>、PSPNet<sup>[16]</sup>、RefineNet<sup>[17]</sup> 和 DeepLab 系列<sup>[18]</sup> 在更多遥感应用中发挥了优势。然而,霍伊斯

(Hoeser) 等人<sup>[12]</sup>指出, 虽然 DCNN 在高分辨率遥感影像场景分类、目标检测和语义分割应用中取得了一定成功, 但多聚焦于城市场景。当前, 由于缺少滨海生态场景深度学习基准数据集和复杂多尺度特征影响, 面向滨海生态监管的研究仍以传统遥感方法为主, 而基于人工智能的滨海生态监管目标提取、分类精度和泛化性仍然有待改善<sup>[12]</sup>。

因此, 研究以围塘养殖、浮筏养殖和红树林三类滨海区域典型生态监管目标为研究对象, 基于 QGIS 构建面向滨海区域监管的 DCNN 多尺度目标语义分割模型, 集成批归一化和空间置弃算法, 改进 UNet 特征融合策略, 尝试改善模型精度和泛化性能, 为滨海区域生态修复、测绘和综合治理提供决策支持。

## 2 研究方法

### 2.1 滨海生态监管多目标语义分割数据集构建

遥感影像具有多波段、大景幅、多时相等特征, 使用面向数码图像的工具标注遥感影像捉襟见肘, 例如读存速度慢、不支持多波段显示等。因此, 研究基于 QGIS 软件的矢量化和矢栅转换工具对 3 类目标地物和背景进行标注。矢量化将遥感影像作为底图, 在创建的面状 Shapefile 格式图层上手动标绘目标对象。其中, 围塘养殖、红树林、浮筏养殖和背景的唯一标识字段分别定义为可区分整型数值。矢栅转换指将标绘完成的 Shapefile 矢量图层转换为栅格图像, 主要采用 QGIS 面转栅格工具完成。为保证矢栅转换后图像大小和像元位置一致, 研究设置了图像输出范围及栅格捕捉等环境变量。经矢栅转换后的文件被保存为 TIFF 格式栅格图像。

研究使用的高分辨率遥感影像源于高分系列卫星产品, 空间分辨率被重采样至 1 米, 像素位深为 8 位。通过提取可见光波段进行真彩色合成后无压缩地保存为 TIFF 格式图像, 用于矢量化和模型训练。由于遥感影像和输出的栅格图像尺寸较大, 在模型训练和预测过程中, 如果直接将其输入到网络模型会造成内存或显存溢出。因此, 研究引入了支持多波段操作、面向 Python 语言的地理空间数据抽象库 (geospatial data abstraction library, GDAL), 对 TIFF 格式的遥

感影像和标注的栅格图像进行裁剪。裁剪图块大小取决于目标尺度和用于模型训练的 CPU 或 GPU 性能。研究综合考虑了目标地物尺度、图像空间分辨率和硬件条件,采用  $256 \times 256$  窗口对遥感影像和标注的栅格图像进行裁剪,并最终获得若干组  $256 \times 256$  像素大小的 TIFF 格式样本图块。一个样本包含一幅输入遥感影像图块和一幅标注栅格图块。图块的文件名保持相同,以便模型正确识别。研究根据机器学习样本数据集划分经验对样本进行随机抽取划分得到训练集、验证集和测试集。

## 2.2 深度卷积神经网络模型 UNet 改进与训练

在除背景类别外的 3 类监管目标地物中,浮筏养殖相较于红树林和围塘养殖的空间尺度较小。为充分利用低级特征图细节信息提高小尺度目标分类精度,研究采用了基于跳跃连接尺度融合策略的 UNet 模型进行实验,并对 UNet 进行了改进:将批归一化 (batch normalization, BN)<sup>[19]</sup> 和空间置弃 (spatial dropout, SPD)<sup>[20]</sup> 模块引入多目标语义分割任务中,改善模型分类精度和泛化能力。SPD 与 Dropout 相比,随机将若干特征图层置弃,不仅实现降维,同时可以减少噪声。Dropout 算法随机地将特征图中独立神经元数值置零,如果相邻特征之间存在较强的相关性,则 Dropout 无法对输出进行归一化处理,而 SPD 有助于提高特征之间的独立性。图 1 展示了 Dropout 和 SPD 对特征图处理的区别。另外,虽然有研究已在 UNet 中引入 BN 层,但多将其置于单个卷积层之后、激活层之前,而本研究将 BN 置于若干个卷积层 (卷积块) 和激活层之后,如图 2 所示。

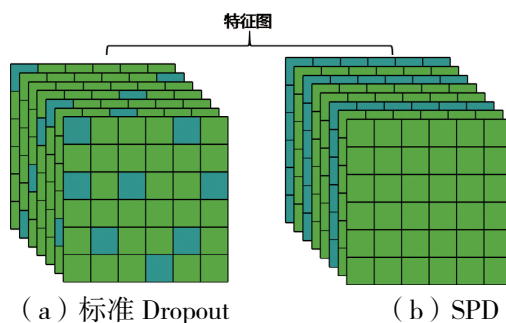


图 1 标准 Dropout 与 SPD 特征图处理

Figure 1 The standard Dropout and SPD feature map processing

研究设计了 4 组实验，实验组分别命名为：MUNet、MUNet-BN、MUNet-BN-SPD 和 MUNet-BN-SPD-500。其中，MUNet 对标准 UNet 结构中卷积核数量和层数进行了修改；MUNet-BN 在 MUNet 结构基础上增加了 BN 层，并将其设置于若干个卷积层和激活层之后；MUNet-BN-SPD 在 MUNet-BN 基础上增加了 SPD 层；MUNet-BN-SPD-500 表示训练迭代次数为 500，而前 3 组设置为 200。

图 2 以 MUNet-BN-SPD 为例展示了研究改进的 UNet 模型网络结构。模型的输入为 3 通道样本图块。其中卷积核采用  $3 \times 3$  像素大小；每个卷积块中卷积核数量分别取值 32、64、128、256 和 512。基于填补处理后，每个卷积块中的各特征图分辨率保持一致。所有经过最大池化操作的特征图尺寸缩小为原始尺寸的一半。每个卷积块最后一个输出层，如 MUNet 的卷积层，MUNet-BN 的 BN 层和 MUNet-BN-SPD 的 SPD 层，采用双线性插值法上采样还原为与前一个卷积块特征图一致空间分辨率的上采样层。特征融合采用跳跃连接，将上采样图层与低层卷积块中相应的卷积层或 BN 层做融合处理。融合方法采用特征图堆砌，即沿着图层通道叠加。最后，使用  $1 \times 1$  卷积操作与 Softmax 分类器获得包含 4 个类别（包含 3 类目标和 1 类背景）的输出分类图。

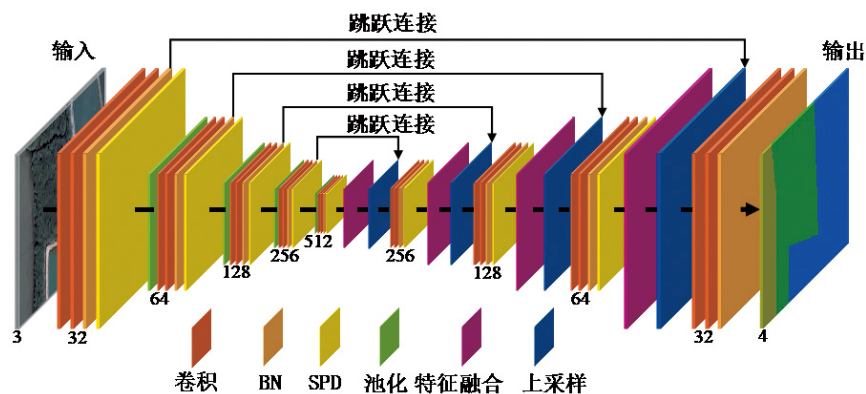


图 2 改进的 UNet 模型网络结构

Figure 2 The improved UNet architecture

DCNN 模型超参数主要包括学习率、样本分组大小和迭代次数，改进的 UNet 模型分别采用 0.001、64 和 200 进行实验。模型优化器、分类器、激活函

数和损失函数分别采用了 Adam<sup>[21]</sup>、Softmax、ReLU 和交叉熵函数。本研究采用的面向多目标分类的交叉熵函数如式(1)所示。

$$L = -\frac{1}{N} \sum_i \sum_{j=1}^M y_{ij} \log(p_{ij}) \quad (1)$$

式中,  $L$  为全部样本的累积损失;  $N$  为样本总数;  $i$  为样本编号;  $M$  为待分类目标数量;  $j$  为目标类别;  $y_{ij}$  为符号函数, 如果样本  $i$  的真实类别为  $j$ , 则  $y_{ij}$  取值为 1, 否则取值为 0;  $p_{ij}$  表示样本  $i$  属于类别  $j$  的概率。

为防止类不平衡引起的评价偏斜问题, 研究除采用总体精度作为评价度量外, 同时还引入了平均交并比 (mean Intersection over Union, mIoU)、频权交并比 (Frequency Weighted Intersection over Union, FWIoU)<sup>[22]</sup> 和 Kappa 系数作为评价指标。其中, mIoU 是计算机视觉中评价语义分割的标准度量, 其计算所有目标类别交集和并集之比的平均值, 如式(2)所示。频权交并比根据每个目标类别出现频率设置权重, 权重乘以每个目标类别的交并比并求和, 如式(3)所示。

$$\text{mIoU} = \frac{1}{n} \sum_{i=0}^n \frac{P_{ii}}{\sum_{j=0}^n P_{ij} + \sum_{j=0}^n P_{ji} - P_{ii}} \quad (2)$$

式中,  $n$  为目标类别数;  $i$  为真实类别;  $j$  为预测类别;  $p_{ij}$  表示将  $i$  预测为  $j$ ;  $p_{ii}$  表示将  $i$  预测为  $i$ ;  $p_{ji}$  表示将  $j$  预测为  $i$ 。

$$\text{FWIoU} = \frac{1}{\sum_{i=0}^n \sum_{j=0}^n p_{ij}} \sum_{i=0}^n \frac{P_{ii}}{\sum_{j=0}^n P_{ij} + \sum_{j=0}^n P_{ji} - P_{ii}} \quad (3)$$

式中,  $n$  为目标类别数;  $i$  为真实类别;  $j$  为预测类别;  $p_{ij}$  表示将  $i$  预测为  $j$ ;  $p_{ii}$  表示将  $i$  预测为  $i$ ;  $p_{ji}$  表示将  $j$  预测为  $i$ 。

另外, 由于遥感影像尺寸较大, 考虑到显存或内存溢出问题, 需要将大图裁剪成较小的子图块输入模型进行预测, 最后将预测结果按照裁剪顺序拼接成与原始影像对应大小和像素位置的分类图像。如果采用常规的滑动窗口裁剪、预测和拼接, 由于每个裁剪样本边缘区域的语义信息不完整, 裁剪边缘处的小尺度对象分类精度不高, 并且容易得到具有明显拼接痕迹的分类图像。因此, 研究采用边缘叠置方法, 即有重叠地裁剪影像并在拼接时采取忽略边缘的策略。如图3所示, 假设实际裁剪图像预测结果为  $B$ , 拼接结果为  $s$ , 如果  $s$  面积  $B$  占面积的百分比为  $\lambda$ , 则相邻裁剪图像的重叠比例将设置为  $1 - \lambda^{1/2}$ 。



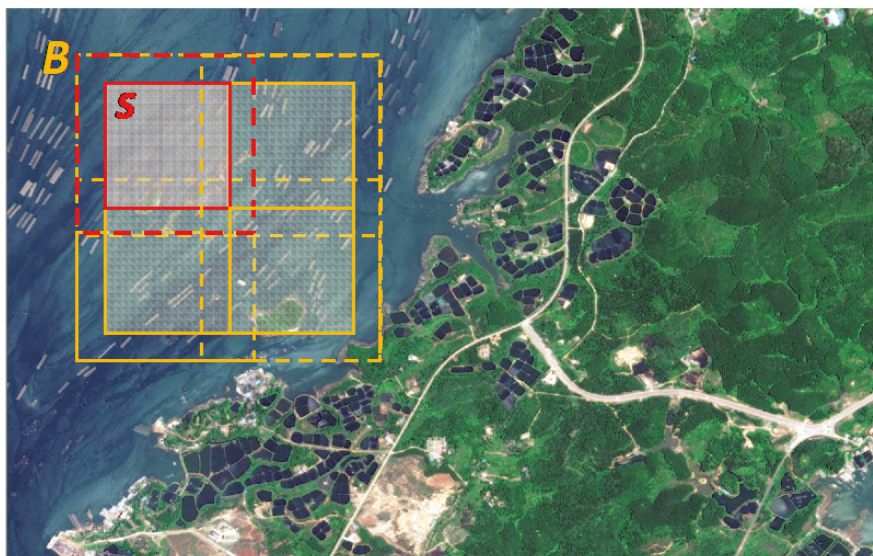


图 3 边缘叠置裁剪

Figure 3 Edge overlay cropping

### 3 结果与讨论

#### 3.1 多目标滨海生态监管语义分割数据集

基于矢量化、矢栅转换和子图裁剪,研究构建了包含 10934 个样本的多目标滨海生态监管语义分割数据集 (multi-object coastal supervision semantic segmentation dataset, MO-CSSSD)。MO-CSSSD 中每个样本分别由大小为  $256 \times 256$  像素的 TIFF 格式遥感图块和标注栅格图块组成。全部样本被划分为训练集、验证集和测试集,其分别包含的样本数量为 8734、1100 和 1100。图 4 展示了数据集样本示例。图中黄色区域表示浮筏养殖,蓝色区域表示围塘养殖,绿色区域表示红树林,而灰色区域表示背景。背景主要涵盖建筑区、海水、裸地和其他植被。需要说明的是,在 MO-CSSSD 中,相邻区域的围塘养殖被标注为整体,不单独分离其岸埂。这些岸埂噪声相较于目标像元较少,同时包含岸埂的围塘养殖构成了独特的模式,有助于识别分类。另外,该噪声可作为基准,用于验证和判定不同模型的泛化能力。

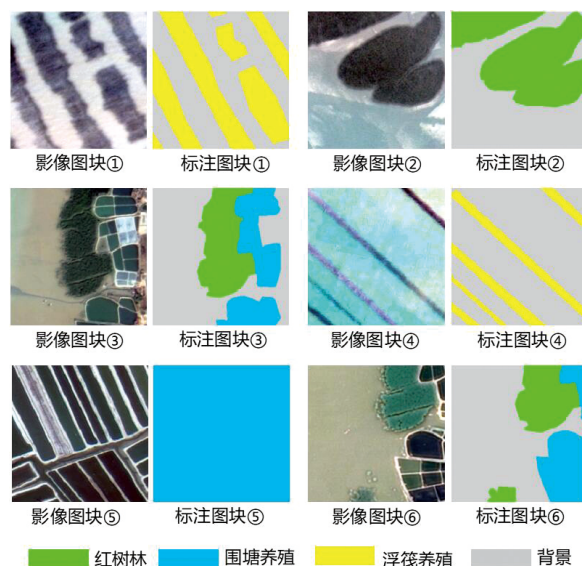


图 4 MO-CSSSD 样本示例

Figure 4 The MO-CSSSD examples

### 3.2 模型测试与评价

改进的 UNet 模型实验结果如图 5 和表 1 所示。从定量评价指标值可知，模型 MUNet-BN 整体表现最佳，其总体精度、mIoU、FWIoU 和 Kappa 系数值最高。而不包含 BN 和 SPD 模块的 MUNet 在本数据集上表现较差，并在训练过程中产生了严重过拟合。

尽管 MUNet-BN 在本数据集上获得了综合最优结果，但从图 5 训练集和验证集变化曲线可以看出，其在验证集上的波动较大，Loss 损失值多高于 0.4，迭代后期有轻微过拟合。相比较而言，MUNet-BN-SPD 虽然在相同迭代次数内各评价指标值低于 MUNet-BN 结果，但差异较小，并且前者在训练过程中的验证集 Loss 损失值更低，与训练集变化一致，拟合得更好。为量化这种差异，研究将实验 1~3 组每轮迭代的训练 Loss 损失值和验证 Loss 损失值做均方误差计算，得到的结果分别是：0.75、0.29 和 0.02。MUNet-BN-SPD 表现出最好的过拟合抑制特性。因此，面对多目标语义分割或分类任务，BN 与 SPD 的联合使用有助于缓解因噪声或类不平衡导致的过拟合。



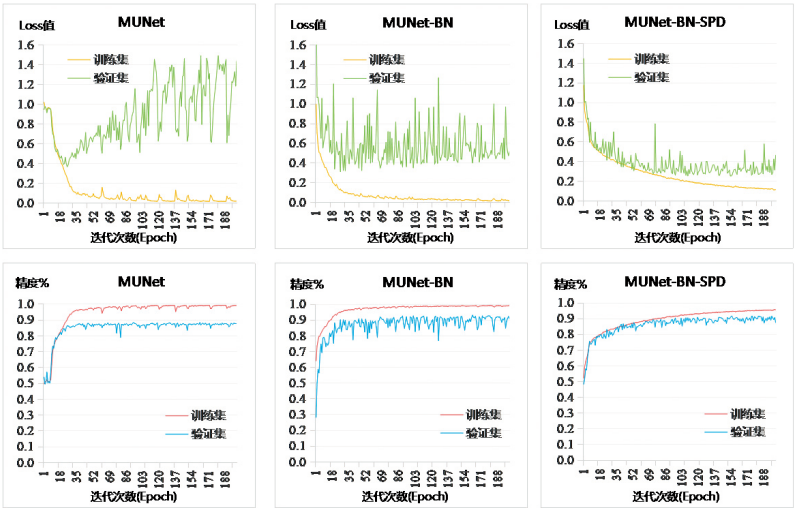


图 5 模型 Loss 损失与精度变化曲线

Figure 5 Loss and accuracy curves of the models

表 1 测试集评价结果

Table 1 Test set evaluation

组号	实验	总体精度 (OA)	平均交并比 (mIoU)	频权交并比 (FWIoU)	卡帕系数 (Kappa)
1	MUNet	0.87	0.68	0.78	0.77
2	MUNet-BN	0.92	0.82	0.86	0.87
3	MUNet-BN-SPD	0.91	0.80	0.84	0.85
4	MUNet-BN-SPD-500	0.92	0.80	0.85	0.86

模型 MUNet-BN-SPD 训练过程中，其精度仍然表现出上升趋势。因此，研究基于原网络结构和超参数设置，增加了迭代轮数，再次训练，并计算了其各项评价指标。结果如表 1 组号 4 所示。随着迭代轮数的增加，MUNet-BN-SPD 的各项评价指标值逼近或等于 MUNet-BN 结果。

图 6 展示了研究选取的具有代表性的测试样本可视化预测结果。其中，影像 1 类别要素包含红树林、围塘养殖及易与围塘养殖混淆的海水背景类；影像 2 包含红树林、围塘养殖及易与红树林混淆的其他植被背景类；影像 3 包含多尺度浮筏养殖和海水背景类；影像 4 包含围塘养殖和复杂背景类（同时包含海水、

裸土和建筑)。另外,由于数据集中未有同时包含4类目标的样本,为验证模型处理所有多目标类别能力,研究从测试集中随机选择了4组样本,经过裁剪、重采样和图像拼接后,构建了一组融合样本,并将其输入模型进行预测。在图6中,样本从左至右依次为:输入遥感图像、真实标注栅格图块、MUNet 分类图、MUNet-BN 分类图、MUNet-BN-SPD 分类图和 MUNet-BN-SPD-500 分类图。

可视化结果更直观地展示了模型对包含不同尺度目标和复杂输入图块的响应细节。结果表明 MUNet 模型在本数据集上的预测误差较大,而增加了 BN 和 SPD 结构的改进模型可以更好地提取分类出目标细节,例如影像1和影像5提取了更完整的红树林。在影像4中,研究设置了噪声,将部分围塘养殖标注为背景类别,用来验证模型泛化能力。结果表明4个模型均可较好地辨别噪声,获得正确分类结果。另外,值得注意的是,对于小尺度浮筏养殖预测分类,增加 SPD 模块并进行特征融合模型比仅采用 BN 结构的模型得到的结果更精细,例如影像3和5。在影像4中,MUNet-BN-SPD 将海水区域识别成围塘养殖,表明上述组合也可能带来冗余噪声,需要进一步优化。

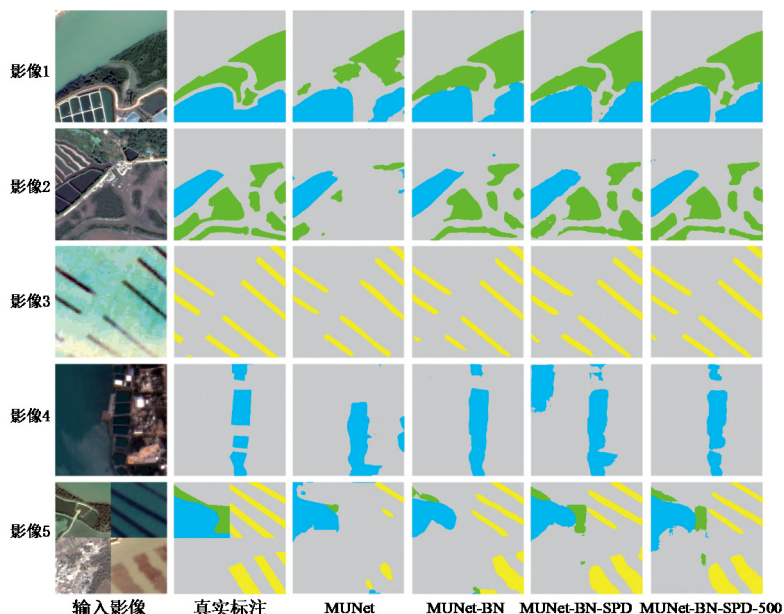


图6 代表性测试样预测结果

Figure 6 Representative test samples prediction

## 4 结论与展望

本研究以围塘养殖、浮筏养殖和红树林 3 类典型滨海生态监管目标为研究对象, 构建了面向滨海生态监管场景的多目标语义分割数据集 MO-CSSSD, 并引入批归一化和空间置弃模块改进 UNet 模型, 改善多目标语义分割任务精度和泛化性能。研究构建的深度学习数据集 MO-CSSSD 为基于人工智能的滨海生态监管目标提取和分类提供研究基础; 置于卷积块和激活层后的批归一化与空间置弃模块组合对多目标语义分割过拟合有明显抑制作用, 促进模型精度和泛化性能的改善; 空间置弃与特征融合组合策略可保存更多目标对象空间细节, 改善目标分类精度, 但也将引入冗余噪声, 需要进一步优化。在未来研究工作中, 作者将基于 MO-CSSSD 数据集, 探索多尺度目标的提取与分类精度的改善, 并进一步优化模型, 减小特征融合引入的冗余噪声影响。

## 参考文献

- [1] 贺义雄, 宋伟鸣, 杨帆. 长三角海洋资源环境一体化治理策略研究: 基于海洋生态系统服务价值影响分析 [J]. 海洋科学, 2021, 45 (6): 63–78.
- [2] 晁明灿, 赵强, 杨铁利, 等. 基于 Landsat8 影像的蓝藻水华提取方法对比研究 [J]. 大气与环境光学学报, 2021, 16 (6): 520–528.
- [3] 周小成, 汪小钦, 向天梁, 等. 基于 ASTER 影像的近海水产养殖信息自动提取方法 [J]. 湿地科学, 2006, 4 (1): 64–68.
- [4] Kang J M, Sui L C, Yang X M, et al. Sea surface-visible aquaculture spatial-temporal distribution remote sensing: a case study in Liaoning province, China from 2000 to 2018 [J]. Sustainability, 2019, 11 (24): 71–86.
- [5] 徐福祥, 高志强, 郑翔宇, 等. 基于 MODIS 数据的 2016 年黄海绿潮灾害动态监测研究 [J]. 海洋科学, 2017, 41 (5): 80–84.
- [6] 初佳兰, 赵冬至, 张丰收, 等. 基于卫星遥感的浮筏养殖监测技术初探: 以长海县为例 [J]. 海洋环境科学, 2008, 27 (S2): 35–40.

- [ 7 ] 胡园园, 范剑超, 王钧. 广义统计区域合并的 SAR 图像浮筏养殖信息提取 [ J ] . 中国图象图形学报, 2017, 22 ( 5 ) : 610–621.
- [ 8 ] Andromachi C, Konstantions T. Impact of intense aquaculture on coastal environments seen by SAR [ C ] // IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2020.
- [ 9 ] Wang M, Cui Q, Wang J, et al. Raft cultivation area extraction from high resolution remote sensing imagery by fusing multi-scale region–line primitive association features [ J ] . ISPRS Journal of Photogrammetry and Remote Sensing, 2017 ( 123 ) : 104–113.
- [ 10 ] Liu Y M, Wang Z H, Yang X M, et al. Satellite-based monitoring and statistics for raft and cage aquaculture in China’s offshore waters [ J ] . International Journal of Applied Earth Observation and Geoinformation, 2020 ( 91 ) : 102118.
- [ 11 ] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [ J ] . Advances in neural information processing systems, 2012 ( 25 ) : 1097–1105.
- [ 12 ] Hoeser T, Bachofer F, Kuenzer C. Object detection and image segmentation with deep learning on Earth observation data: A review—Part II: Applications [ J ] . Remote Sensing, 2020, 12 ( 18 ) : 3053.
- [ 13 ] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [ C ] //Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431–3440.
- [ 14 ] Sharma A, Liu X W, Yang X J, et al. A patch-based convolutional neural network for remote sensing image classification [ J ] . Neural Networks, 2017 ( 95 ) : 19–28.
- [ 15 ] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation [ C ] //International Conference on Medical image computing and computer-assisted intervention. Springer Cham, 2015:

234–241.

- [ 16 ] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network [ C ] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2881–2890.
- [ 17 ] Lin G S, Milan A, Shen C H, et al. RefineNet: multi-path refinement networks for high-resolution semantic segmentation [ C ] //In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition ( CVPR ). IEEE, 2017.
- [ 18 ] 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, 2016 ( 40 ) : 834–848.
- [ 19 ] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [ C ] //International conference on machine learning. PMLR, 2015: 448–456.
- [ 20 ] Tompson J, Goroshin R, Jain A, et al. Efficient object localization using convolutional networks [ C ] //Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 648–656.
- [ 21 ] Kingma D P, Ba J. Adam: A method for stochastic optimization [ J ] . arXiv e-prints, 2014.
- [ 22 ] Garcia G A, Orts S, Oprea S, et al. A review on deep learning techniques applied to semantic segmentation [ J ] . arXiv e-prints, 2017.



## Multi-scale Objectives Semantic Segmentation for Coastal Ecological Supervision

Chen Yan<sup>1</sup> Yang Xiaotong<sup>2</sup> Xi Yantao<sup>3</sup> Xu Lixiang<sup>1</sup> Li Xinlu<sup>1</sup>

*1. School of Artificial Intelligence and Big Data, Hefei University, Hefei;*

*2. Department of Marine Mapping and Geographic Information, National Marine Data and Information Service, Tianjin;*

*3. School of Resources and Geosciences, China University of Mining and Technology, Xuzhou*

**Abstract:** To improve the lack of deep learning dataset of coastal ecological scenes and low accuracy of multi-scale objectives semantic segmentation for remote sensing image classification, we take three types of coastal typical ecological supervision multi-scale objectives of mangrove, raft cultivation and pond aquaculture as research objects, constructs a benchmark dataset for coastal ecological supervision, improves the UNet feature fusion by integrating batch normalization and spatial dropout modules, and proposes a multi-scale deep convolutional semantic segmentation model. The model has an overall accuracy of 92% on the test set, a kappa coefficient of 0.87, and a mIoU of 82%. The experimental results show that the coupled stacking of batch normalization and feature fusion spatial dropout can effectively suppress multi-scale objectives semantic segmentation overfitting and improve the model accuracy and generalization performance. The proposed model and the constructed semantic segmentation dataset for coastal ecological supervision can provide decision support for ecological restoration, mapping and comprehensive management in coastal areas.

**Key words:** Coastal ecological supervision; Artificial intelligence; Mangroves; Mariculture; Benchmark dataset; Multi-scale feature fusion; Semantic segmentation