

基于多尺度特征模糊卷积神经网络的遥感图像分割

马翔悦¹,徐金东¹,倪梦莹^{2*}

(1.烟台大学计算机与控制工程学院,山东烟台 264005; 2.烟台大学物理与电子信息学院,山东烟台 264005)

摘要:为解决高分辨率遥感图像“同谱异物、同物异谱”的不确定性以及大量空间信息利用率低的问题,提出一种基于多尺度特征的模糊卷积神经网络模型。该模型在长跳跃连接部分加入模糊学习模块去除噪声特征,缓解类别间的不确定性;利用多孔空间金字塔池化融合多尺度特征,提取完备的空间上下文信息,提升分割性能。试验结果表明,该模型在 Potsdam 数据集和 Vaihingen 数据集上的整体准确度分别达到 92.65% 和 93.19%, 明显优于现有流行的深度学习模型,能够显著提升高分辨率遥感图像的语义分割性能。

关键词:模糊学习;多孔空间金字塔池化;多尺度特征;编码器-解码器;卷积神经网络

中图分类号:TP751 **文献标志码:**A

引用格式:马翔悦,徐金东,倪梦莹.基于多尺度特征模糊卷积神经网络的遥感图像分割[J].山东大学学报(工学版),2024,54(3):44-54.

MA Xiangyue, XU Jindong, NI Mengying. Remote sensing image segmentation based on multi-scale feature fuzzy convolutional neural network[J]. Journal of Shandong University (Engineering Science), 2024, 54(3):44-54.

Remote sensing image segmentation based on multi-scale feature fuzzy convolutional neural network

MA Xiangyue¹, XU Jindong¹, NI Mengying^{2*}

(1. School of Computer and Control Engineering, Yantai University, Yantai 264005, Shandong, China; 2. School of Physics and Electronic Information, Yantai University, Yantai 264005, Shandong, China)

Abstract: In order to solve the uncertainty of "same spectrum of different objects, the same object with different spectrum" and the low utilization rate of large amount of spatial information in high-resolution remote sensing images, a fuzzy convolutional neural network model based on multi-scale features was proposed. The fuzzy learning module was added to the long jump connection to remove noise features and ease the uncertainty between classes. The multi-scale features were fused by atrous spatial pyramid pooling, and complete spatial context information was extracted to improve segmentation performance. The experimental results showed that the overall accuracy of the model on Potsdam dataset and Vaihingen dataset reached 92.65% and 93.19%, respectively, which were significantly better than the existing popular deep learning models and could significantly improve the semantic segmentation performance of high-resolution remote sensing images.

Keywords: fuzzy learning; atrous spatial pyramid pooling; multi-scale features; encoder-decoder; convolutional neural networks

0 引言

随着高分辨率遥感技术的飞速发展^[1],高分辨率遥感图像(high-resolution remote sensing image, HRI)语义分割在国土资源利用管理^[2]、环境保护监

督^[3]、城市规划^[4]等诸多领域都具有重要研究意义。在 HRI 中相同的地物应由同种物质组成且具有相似的特征,但由于成像过程中受噪声、大气辐射等诸多因素的影响,导致 HRI 往往存在类内异质性、类间相似性的不确定性现象^[5]。因此,如何充分考虑图像中丰富的上下文信息、精确分割 HRI 是

长期以来需要面对的重要问题。

语义分割是图像分割过程中为每个像素分配唯一的标记,具有相同标记的像素具有某种共同的视觉特性,使图像更容易被理解和分析^[6],目前HRI的对象边界与小物体的错误分类问题是影响分割精度的主要因素。基于图像的颜色、纹理、灰度等底层特征的阈值分割^[7]和聚类分割^[8]等方法难以应对HRI中地物复杂性。近年来,随着深度学习的发展,编码器-解码器结构^[9]在图像分割领域广受青睐,出现了众多优秀的深度学习模型,能够同时提取HRI的浅层和深层特征^[10]。文献^[11]提出了全卷积神经网络(fully convolutional networks, FCN),采用上采样卷积操作代替卷积神经网络的全连接层生成密集像素,但多池化操作会损失原始特征图中的信息,降低特征图的分辨率,且小目标信息不完整^[12]。DeepLab中的多孔空间金字塔池化结构(atrous spatial pyramid pooling, ASPP)利用多个不同卷积核的空洞卷积分支,从多尺度角度分析和提取对象特征^[13-14],展现了更好的HRI分割性能。但仅使用双线性插值恢复特征图,难以捕捉对象的精细边界。文献^[15]结合DeepLab v3+网络和基本对象的图像分析方法提高HRI语义分割性能。文献^[16]通过构建邻域像素亲和损失函数,使模型更精准分割遥感图像的目标边界。文献^[17]提出STLNet,引入纹理增强模块捕捉与纹理相关的信息并增强纹理细节,引入金字塔纹理特征提取模块提取多尺度统计特征信息,STLNet分析了网络模型底层次信息分布,并进行有效的语义分割。

虽然基于深度学习的遥感图像分割方法在HRI中取得了一定成功,但尚存在以下缺陷:没有针对性地考虑HRI中类内异质性、类间相似性的不确定性问题;网络浅层特征不足,获取的上下文信息不完整,梯度易消失。因此,本研究提出一种全新的基于多尺度特征的模糊卷积神经网络(fuzzy convolutional neural network based on multiscale features, MF-FCNN)用于HRI语义分割,主要内容如下:(1)针对HRI分割中受噪声影响的不确定性问题,设计了模糊学习模块(fuzzy learning module, FLM),采用高斯形式的隶属度约束多个参数描述,为每个特征点分配相应的模糊语义标记,并融入网络的长跳跃连接部分。(2)考虑网络浅层特征不足、获取上下文信息不完整的问题,在网络模型中聚合改进的ASPP结构,扩大感受野并充分利用空间上下文信息,捕捉小目标特征。(3)采用完全残差模块代替原始残差模块降低网络梯度消失

的风险,同时,为增强分割目标的学习过程,根据焦点损失和HRI中小目标分割问题,改进并提出了新的损失函数。

1 相关工作

卷积神经网络将输入图像的位置信息变化映射到输出中,从而实现提取图像中局部形状、边缘分布等高层抽象特征的功能。常见的卷积神经网络包含输入层、输出层、卷积层、池化层、激活函数层、批量归一化层(batch normalization, BN)等。卷积神经网络的相关结构介绍如下。

(1) 卷积层。卷积层是卷积神经网络的核心,以一组固定大小的滤波器逐一对图像中的区域做内积操作并将结果传递至下一层。其中,卷积神经网络的浅层卷积核提取边缘特征信息,深层卷积核提取高层抽象语义信息。

(2) 池化层。池化又称为降采样操作。本研究采用最大池化层减小特征图尺寸,降低计算量。另外,池化层可以使其后卷积层中的卷积核对应原始图像区域中更大的感受野,有效保留并提取特征图的纹理信息。

(3) 激活函数层。本研究采用ReLU激活函数,ReLU激活函数可以加快网络收敛速度,降低计算复杂度。

(4) BN层。在每个卷积层后加入BN层,加快模型收敛。

2 方法与原理

MF-FCNN模型结构如图1所示。采用编码器-解码器的架构,编码器由4组卷积层、4个池化层(池化操作采用最大池化,保存最大池化的索引位置)、3个完全残差学习模块以及ASPP组成。完全残差学习模块融合了原始输入信息,可以更好改善因网络层数加深而造成的梯度消失问题,同时弥补因卷积操作造成的特征信息损失,进而增强特征融合,通过更深层实现更好的分割精度^[17]。解码器包含4组上采样与卷积模块,输入特征信息包括自编码器后经过FLM的浅层特征信息、经过ASPP的多尺度特征信息和经过上一层上采样操作的深层信息,各级信息进行融合,还原特征图细节,降低噪声干扰并保持相应的空间信息维度。解码特征输入至 1×1 的卷积层,经损失函数约束获得最终分割结果图。由于增加神经网络深度可能会降低模型精度,因此本研究使用3个完全残差学习模块保持模

型的轻量级,降低模型计算负担。另外,完全残差学习模块和改进的ASPP中的融合方法都是采用简

单的加法运算,这样做尽可能减少了模型的计算复杂度又保持了特征信息的一致性。

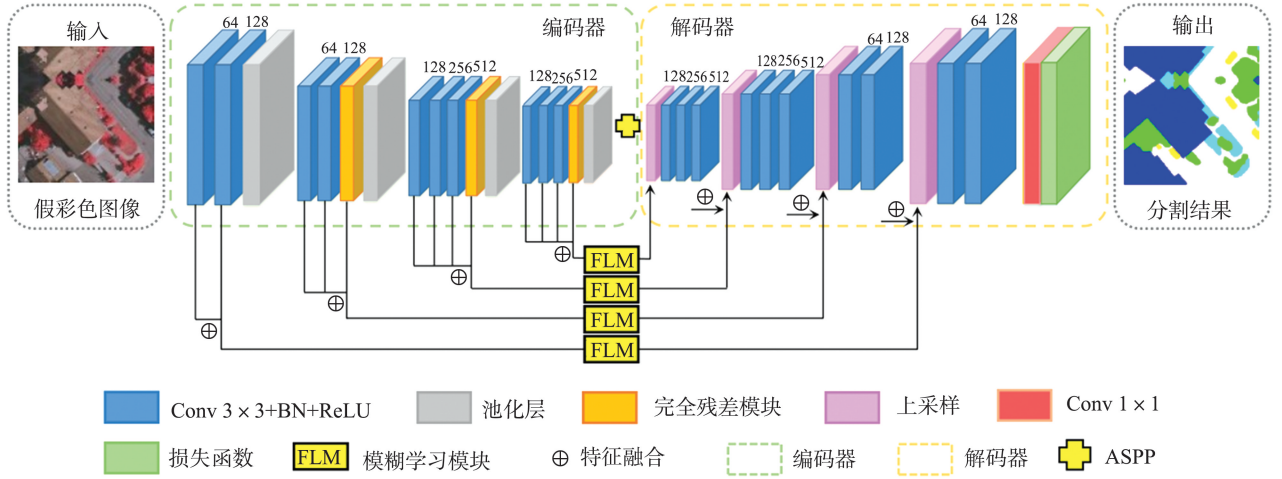


图1 MF-FCNN 结构
Fig.1 Structure of MF-FCNN

2.1 模糊学习模块

在 HRI 中,小目标和边界像素受到混合像元和噪声的影响,导致分割结果存在较多的不确定性。相比传统的精确数理统计方法,模糊理论能够更有效处理不确定性问题,本研究设计并提出了一种新的 FLM,旨在提升 HRI 小目标和边界区域的分割精度。作为网络的一部分,FLM 可以看作由高斯隶属度函数中的多个参数描述的模糊层。FLM 结构如图 2 所示, $H \times W \times C$ 表示输入特征图的大小, H 、 W 、 C 分别为特征图的高度、宽度以及输入特征图的通道数。

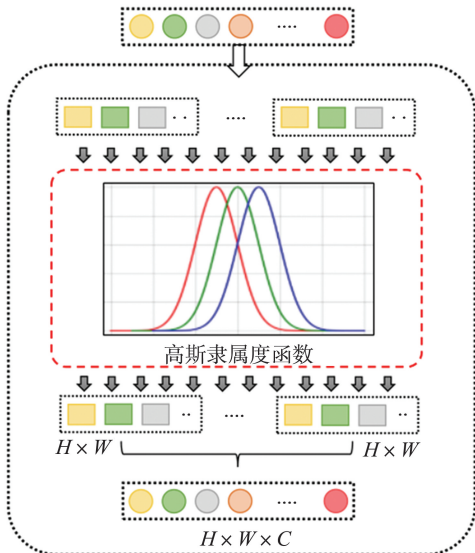


图2 FLM 结构
Fig.2 FLM structure

在 FLM 中,每个高斯隶属度函数为特征点分配一个模糊语义标记,高斯隶属度函数表示为:

$$G_g^c(n_1, n_2) = e^{-\frac{n_1^2 + n_2^2}{2\sigma^2}}, \quad (1)$$

式中: g 和 σ 为常量; c 为通道; n_1 和 n_2 分别为当前像素点 x 和 y 坐标与中心点坐标之间的差值; ∂_1 和 ∂_2 为图像中所选中心点的坐标,即 $n_1 = x_1 - \partial_1, n_2 = x_2 - \partial_2$ 。为所选区域中的所有像素点确定中心点,即:

$$\partial_1 = \frac{1}{i} \times \sum_0^i x_1 + x_2 + x_3 + \dots + x_i, \quad (2)$$

$$\partial_2 = \frac{1}{i} \times \sum_0^i y_1 + y_2 + y_3 + \dots + y_i, \quad (3)$$

式中 i 为所选区域中的像素数。

通道中每个相邻像素点的权重

$$W^c(n_1, n_2) = \frac{G_g^c(n_1, n_2)}{\sum \sum G_g^c}. \quad (4)$$

将每个位置的像素坐标值与权重矩阵中的相应元素相乘,设置中心点在原点,确定 σ 。最后将质心与相邻点的坐标值相加,得到模糊学习后的质心像素值。模糊半径 r 的相关公式为:

$$\tilde{w}_{ij} = e^{-\frac{(x_{ij} - \partial_1)^2 + (y_{ij} - \partial_2)^2}{2\sigma^2}}, \quad (5)$$

$$w_{ij} = \frac{\tilde{w}_{ij}}{\omega}, \quad (6)$$

$$\omega = \begin{cases} \text{Sum}(\tilde{W}), & \sqrt{(x_{ij} - \partial_1)^2 + (y_{ij} - \partial_2)^2} \leq r \\ \omega, & \sqrt{(x_{ij} - \partial_1)^2 + (y_{ij} - \partial_2)^2} \geq r \end{cases}, \quad (7)$$

式中, \tilde{w}_{ij} 和 w_{ij} 分别为初始权重矩阵 \tilde{W} 和最终权重矩阵 W 中的各个元素,Sum(\cdot) 为对 \tilde{W} 中元素的求和。

对图像中的所有点重复此过程,利用 FLM 学习特征图和相应分割结果之间的内在规则,提取模糊

语义特征。

2.2 改进的 ASPP

与自然图像相比, HRI 中地理实体的分布分散而复杂, 很难识别并准确分割小物体。获得较远空间上下文信息需要更大的感受野, 而具有大感受野的 Deeplab-v2 中的 ASPP 采用了多孔卷积, 部分像素点信息易被忽略形成网格效应^[18], 不利于小目标和边界信息获取。因此, 本研究添加了 1×1 卷积克服此类问题, 并附加最大池化分支以合并全局上下文信息。改进的 ASPP 结构图如图 3 所示。

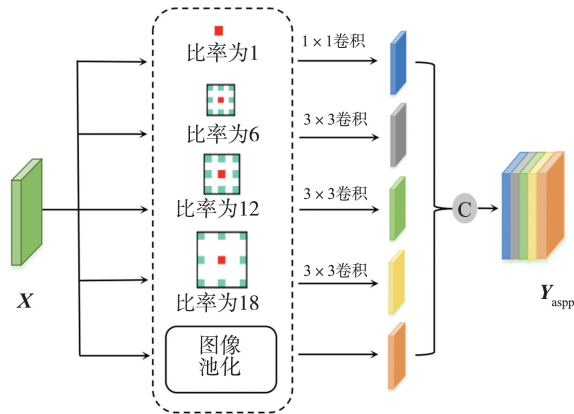


图3 改进的 ASPP 结构

Fig.3 Improved ASPP structure

改进的 ASPP 通过扩张卷积增加感受野获得局部特征, 在保持特征图分辨率大小不变的前提下增大感受野, 即:

$$Y = \text{Concat}(I_{\text{pooling}}(X), H_{1,3}(X), H_{6,3}(X), H_{12,3}(X), H_{18,3}(X)), \quad (8)$$

式中: Y 为输出特征图; X 为输入特征图; $\text{Concat}(\cdot)$ 为对特征图做第一维度上的拼接操作; I_{pooling} 为输入特征图的最大池化特征; $H_{l,k}(X)$ 为带孔卷积, 其中 l 为采样率, k 为卷积核大小。

2.3 损失函数

传统的交叉熵损失是将输出值用 Softmax 函数激活, 并与标记做带权重的交叉熵损失, 若不调节交叉熵损失项权重, 模型会偏向于更新背景相关的权重, 导致难以处理 HRI 中类别不平衡问题。文献 [18] 提出了焦点损失缓解图像中类别不平衡的问题, 即:

$$L_f = - \sum_{(x,y) \in R_i} (1 - P_{x,y})^\alpha \times \log P_{x,y} - \sum_{(x,y) \in R_{\text{non-}i}} (P_{x,y})^\beta \times \log(1 - P_{x,y}), \quad (9)$$

式中: L_f 为焦点损失函数; $P_{x,y}$ 为坐标 (x, y) 的像素级估计目标分类概率; R_i 和 $R_{\text{non-}i}$ 分别为属于目标类和非目标类的坐标集; α 和 β 为具有正值的超参数, 它们决定了损失函数的形状和轮廓。通过将目标

像素与因子 $(1 - P_{x,y})^\beta$ 相乘, 图像中目标类概率大、易于分类的像素贡献减弱, 因此, 在每次迭代中都会强调分类困难的样本, 即图像中目标类概率小、不易于分类的像素贡献增强, 从而缓解分类不平衡的问题。

根据 HRI 分割任务的特点, 本研究提出的损失函数

$$L_{rs} = - \sum_{(x,y) \in R_i} (1 + e^{\alpha + (1 - P_{x,y})^\beta}) \times \log P_{x,y} - \sum_{(x,y) \in R_{\text{non-}i}} (1 + e^{\alpha + (P_{x,y})^\beta}) \times \log(1 - P_{x,y}). \quad (10)$$

设计的损失函数基于以下事实: (1) 在 HRI 中, 道路、建筑物等目标在优化阶段开始时具有非常高的目标类概率, 汽车、单棵树木等小目标所占目标类概率较低。因此, 错误分类这些样本应受到高的惩罚, 且网络会被迫记住此类小目标样本的特征。(2) 目标像素分类概率在早期的优化阶段变化更大, 为目标概率大的像素提供相对较高的损失函数值, 将有助于提高网络优化过程中的鲁棒性和稳定性。

与 L_f 相比, L_{rs} 为图像中目标概率大的像素分配了更高的惩罚因子, 在优化阶段更快纠正小目标分类, 更有利于处理 HRI 中类别分类不平衡问题。

3 试验与分析

3.1 试验设置与数据集

为了定量的评估性能, 试验采用交并比 I_{ou} ^[19]、 F_1 值^[20]、精确率^[21]、召回率^[22]、整体准确率 O_A ^[23] 和平均 I_{ou} ^[24] 作为评价指标。选取了 FCN^[11]、U-Net^[25]、SegNet^[26]、EaNet^[27] 和 MaNet^[28] 深度学习网络作为试验对比分割模型。本研究在 ISPRS^[29] 发布的两个公开 HRI 数据集上进行了方法评估 (数据集获取链接: https://blog.csdn.net/weixin_43451806/article/details/127228501?spm=1001.2014.3001.5502)。

3.1.1 Potsdam 数据集

Potsdam 数据集拍摄于德国 Potsdam 城市。该数据集包含 38 张尺寸相同的全局图像, 图像大小为 6 000 像素 \times 6 000 像素, 分辨率为 0.05 m, 所有图像都具有对应的真值图。数据集包括建筑物、低矮植被、路、汽车、树木、杂物 6 种地物类型。对于 Potsdam 数据集, 训练集选取了 1 282 张局部图像, 测试集选择了 600 张局部图像和 18 张全局图像。

3.1.2 Vaihingen 数据集

Vaihingen 数据集拍摄于德国 Vaihingen 城市。

该数据集包含 33 张尺寸不等的全局图像,数据集包括建筑物、低矮植被、路、汽车、树木、杂物 6 种地物类型。其中汽车仅占标签像素总数的 1.21%,建筑物和路占 50%,因此,汽车等小型物体分割的准确性更能考验模型的分割性能。试验选取了 1 324 张局部图像作为训练集,600 张局部图像和 17 张全局图像作为测试集。

3.1.3 数据集预处理

试验训练网络采用 512 像素×512 像素局部图像。由于原始 Potsdam 数据集、Vaihingen 数据集中全局图像尺寸不等,在试验之前需先对数据集进行预处理操作。本研究使用步长为 128、大小为 512 像素×512 像素的滑动窗口在全局图像与其对应的真值图上进行截取,之后对裁剪后的所有局部图像块分别进

行 90°、180°、270°旋转,水平、垂直、镜像翻转,从而实现数据增广,得到试验所需的训练集和测试集。

3.1.4 试验参数设置

试验采用 Linux 操作系统,建模过程以 PyTorch 框架为基础,程序语言为 Python,硬件配置为 Intel (R) Core (TM) i7-7820X,最大提升频率为 3.6 GHz,一条指令可以执行 4 个单精度浮点运算。在训练过程中,设置动量值为 0.99,权重衰减为 2×10^{-5} ,权重使用标准差为 $\sqrt{2/n}$ 的零均值高斯分布进行初始化(n 为输入个数),训练批尺寸为 4,迭代次数为 1 000。初始学习率为 0.000 1。

3.2 Potsdam 数据集试验结果与分析

Potsdam 数据集的局部图像分割结果如图 4 所示。

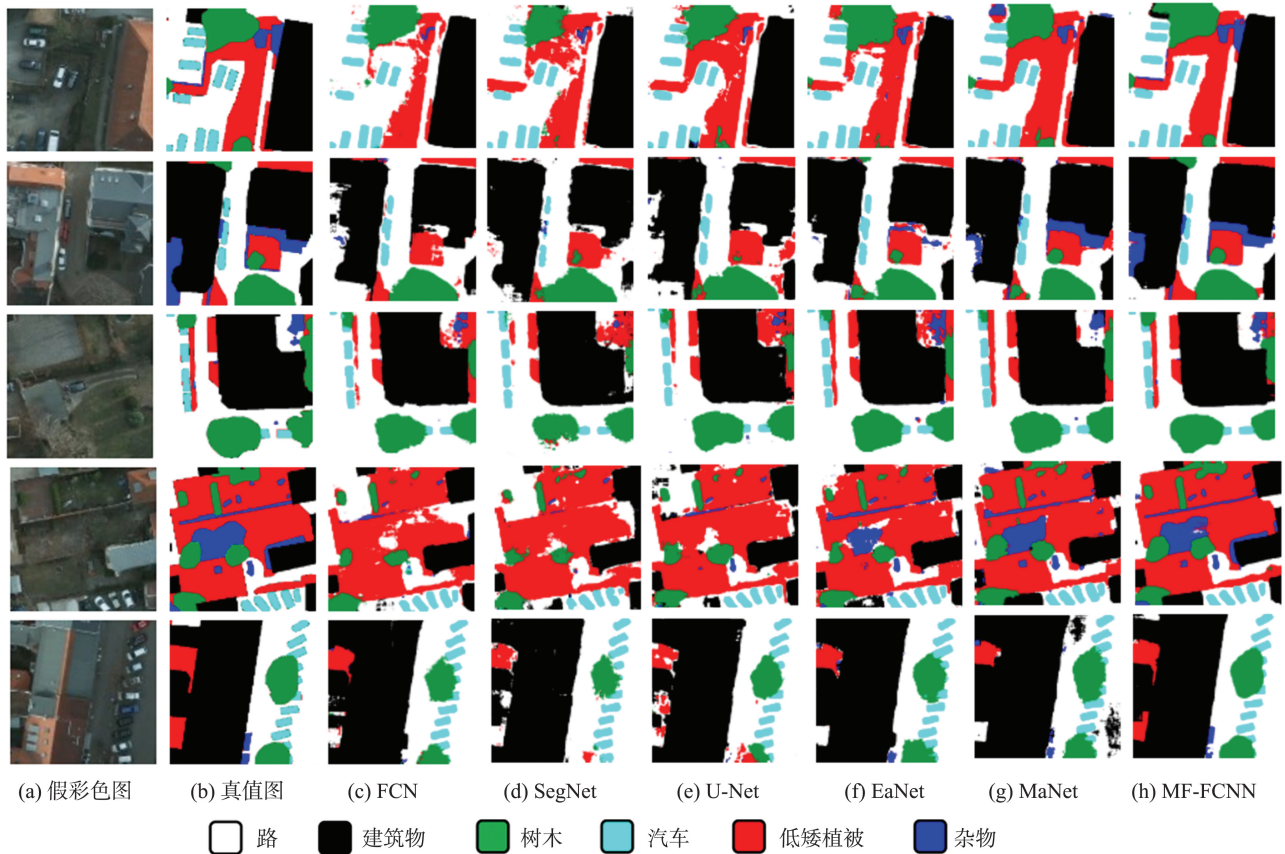


图 4 Potsdam 数据集(局部)的试验结果

Fig.4 Experimental results on the Potsdam dataset (local)

如图 4 所示,FCN 分割结果较差,建筑物边缘分割细节不清晰且内部存在分割不连续的现象。相比于 FCN,SegNet 低矮植被与路混淆分割现象减少,更易于保存目标的细节信息。U-Net 在单棵树木和小汽车的分割上相对于前者更加精细。EaNet 与 MaNet 分别采用金字塔池化和注意力机制捕获丰富多尺度上下文信息,整体分割效果较前几种模型有了很大提升。MF-FCNN 分割细节信息最为完善,尤其对杂物的分割取得了显著的

分割效果。对汽车和单棵树木这类小物体,MF-FCNN 的分割边缘更加光滑,空间一致性较其他对比模型更胜一筹。

为了证明所提出的模型对 HRI 的普适性,充分利用 HRI 中的全局场景信息,本研究对 Potsdam 数据集 6 000 像素×6 000 像素的全局图直接分割,试验结果如图 5 所示。比较可知 MF-FCNN 的分割边缘更加平滑,小物体分割最为精确,试验结果最接近真值图。

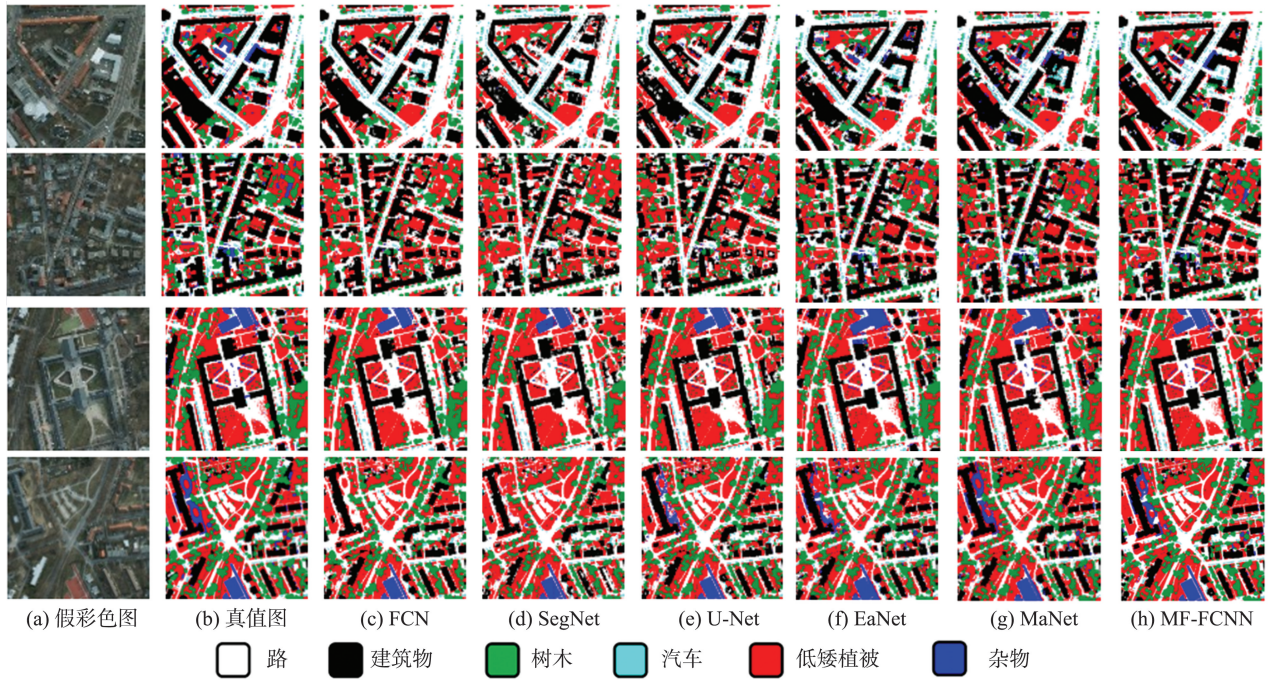


图 5 Potsdam 数据集(全局)的试验结果

Fig.5 Experimental results on the Potsdam dataset (global)

定量评价如表 1~5 所示,表 1~4 分别为 准确率、召回率指标对比结果,表 5 为 Potsdam 数据集 Potsdam 数据集上不同模型的每一类别的 I_{ou} 、 F_1 、精 的平均 I_{ou} 和 O_A 指标计算结果。

表 1 Potsdam 数据集上的 I_{ou} 对比

Table 1 I_{ou} comparison on the Potsdam dataset

模型	I_{ou}					
	路	建筑物	低矮植物	树	汽车	杂物
FCN	0.73	0.87	0.60	0.62	0.71	0.32
SegNet	0.74	0.79	0.72	0.61	0.66	0.49
U-Net	0.73	0.87	0.65	0.65	0.72	0.51
EaNet	0.81	0.93	0.70	0.72	0.76	0.61
MaNet	0.84	0.90	0.78	0.74	0.79	0.60
MF-FCNN	0.90	0.90	0.80	0.85	0.82	0.63

表 2 Potsdam 数据集上的 F_1 对比

Table 2 F_1 comparison on the Potsdam dataset

模型	F_1					
	路	建筑物	低矮植物	树	汽车	杂物
FCN	0.85	0.93	0.75	0.76	0.83	0.48
SegNet	0.85	0.88	0.84	0.76	0.88	0.66
U-Net	0.84	0.93	0.79	0.79	0.83	0.68
EaNet	0.90	0.96	0.82	0.84	0.86	0.80
MaNet	0.92	0.94	0.89	0.87	0.88	0.85
MF-FCNN	0.95	0.95	0.89	0.92	0.89	0.87

表 3 Potsdam 数据集上的精确率对比

Table 3 Precision comparison on the Potsdam dataset

模型	精确率					
	路	建筑物	低矮植物	树	汽车	杂物
FCN	0.77	0.93	0.71	0.86	0.85	0.98
SegNet	0.78	0.89	0.79	0.92	0.86	0.97
U-Net	0.79	0.94	0.74	0.87	0.90	0.94
EaNet	0.87	0.97	0.80	0.89	0.89	0.86
MaNet	0.92	0.92	0.85	0.93	0.89	0.89
MF-FCNN	0.94	0.93	0.88	0.95	0.92	0.95

表4 Potsdam数据集上的召回率对比
Table 4 Recall comparison on the Potsdam dataset

模型	召回率					
	路	建筑物	低矮植物	树	汽车	杂物
FCN	0.94	0.93	0.79	0.68	0.81	0.32
SegNet	0.93	0.87	0.89	0.65	0.74	0.50
U-Net	0.90	0.92	0.85	0.72	0.78	0.53
EaNet	0.93	0.96	0.84	0.79	0.84	0.75
MaNet	0.92	0.97	0.93	0.81	0.86	0.81
MF-FCNN	0.95	0.97	0.90	0.89	0.88	0.80

表5 Potsdam数据集上的平均 I_{ou} 和 O_A 对比
Table 5 Average I_{ou} and O_A comparison on the Potsdam dataset

比较内容	平均 I_{ou}	$O_A/\%$
FCN	0.64	82.79
SegNet	0.69	82.78
U-Net	0.68	84.60
EaNet	0.76	89.18
MaNet	0.78	89.87
MF-FCNN	0.81	92.65

由表1~5可知, MF-FCNN模型的每一类别的度量指标几乎都达到了最高, 尤其对于汽车这种小物体的分割, MF-FCNN具有最高的评价指标。由表5可知, MF-FCNN的平均 I_{ou} 比FCN提高了0.17, O_A 达到92.65%。

3.3 Vaihingen数据集试验结果与分析

Vaihingen数据集局部图像的分割结果如图6所示。

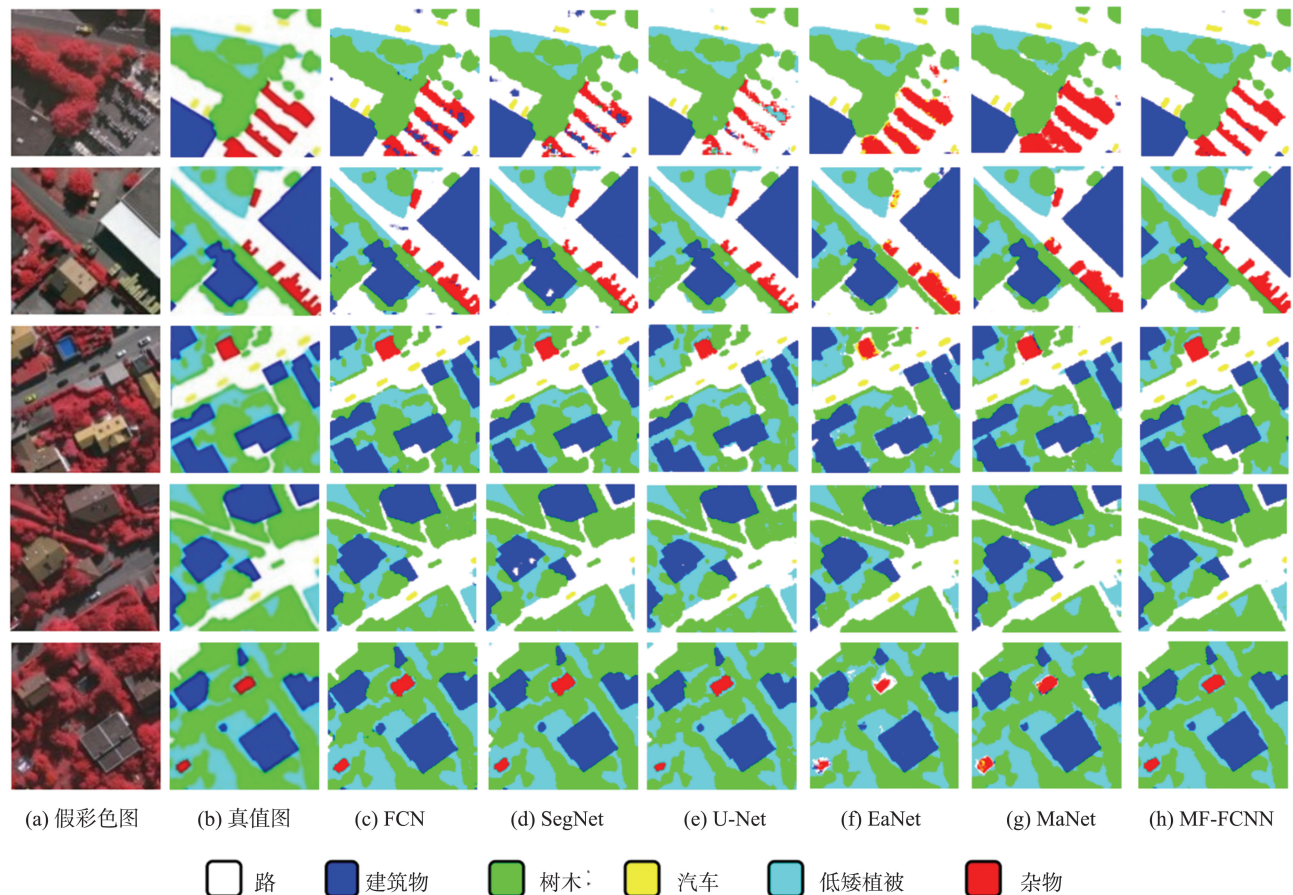


图6 Vaihingen数据集(局部)的试验结果

Fig.6 Experimental results on the Vaihingen dataset (local)

由图6可知: FCN分割效果最差, 没有真实还原建筑物几何特征, 道路分割边缘模糊; SegNet分割效果有一定改善, 能比较准确分割出道路, 但建筑物内部分割仍存在区域信息缺失问题; U-Net能

较精确区分道路、建筑物、低矮植被以及树木, 但针对小汽车和杂物这些在图像中所占比例较少的目标, 其细节显示性能较差, 杂物和汽车边缘存在锯齿状; EaNet整体分割效果较好, 但目标边缘分割粗

糙,对杂物的分割效果不明显。MaNet 相较于前面其他模型分割结果最为精确,但对于杂物的分割仍旧存在误分类现象。MF-FCNN 模型细节表现性能最为优越,能精准捕捉到分割目标的形状和区域位置信息,特别是杂物的分割,与其他上述模型相比

表现尤为突出。

为了考虑全局场景信息,采用 Vaihingen 数据集的全局图作为直接预测,试验结果如图 7 所示。可视化试验结果证明了 MF-FCNN 模型的有效性和可靠性。

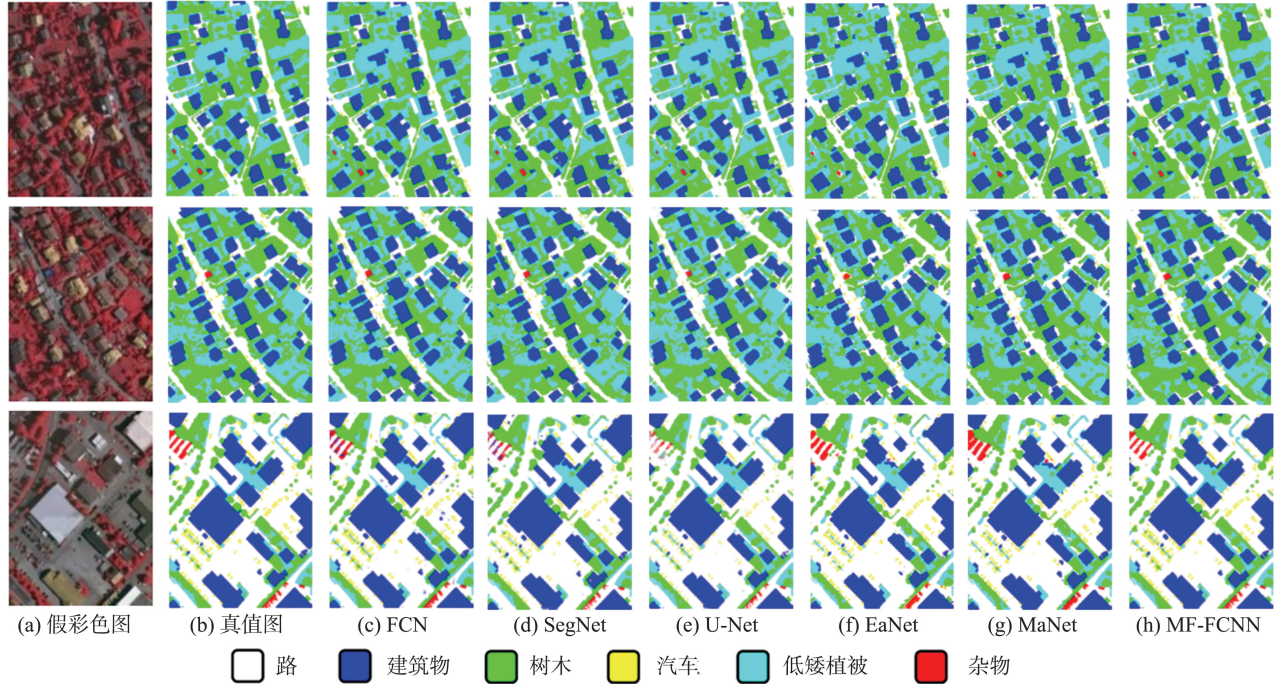


图7 Vaihingen 数据集(全局)的试验结果

Fig.7 Experimental results on the Vaihingen dataset (global)

Vaihingen 数据集中各模型的 I_{ou} 、 F_1 、精确率、召回率如表 6~9 所示,整体来看,MF-FCNN 的综合得分最高,分割性能最好。表 10 给出了各深度学习

模型在 Vaihingen 数据集上的平均 I_{ou} 和 O_A 的值。其中,本研究所提模型的平均 I_{ou} 与 O_A 分别达到了 0.81 和 93.19%。

表6 Vaihingen 数据集上的 I_{ou} 对比
Table 6 I_{ou} comparison on the Vaihingen dataset

模型	I_{ou}					
	路	建筑物	低矮植物	树	汽车	杂物
FCN	0.82	0.82	0.69	0.64	0.66	0.43
SegNet	0.81	0.73	0.77	0.69	0.60	0.48
U-Net	0.89	0.88	0.77	0.79	0.75	0.51
EaNet	0.88	0.92	0.72	0.81	0.71	0.64
MaNet	0.90	0.92	0.76	0.83	0.71	0.70
MF-FCNN	0.91	0.93	0.79	0.86	0.67	0.72

表7 Vaihingen 数据集上的 F_1 对比
Table 7 F_1 comparison on the Vaihingen dataset

模型	F_1					
	路	建筑物	低矮植物	树	汽车	杂物
FCN	0.90	0.90	0.82	0.78	0.81	0.57
SegNet	0.90	0.85	0.87	0.84	0.75	0.60
U-Net	0.94	0.94	0.87	0.88	0.86	0.73
EaNet	0.94	0.96	0.84	0.90	0.83	0.77
MaNet	0.95	0.96	0.87	0.91	0.83	0.82
MF-FCNN	0.95	0.96	0.89	0.92	0.90	0.91

表8 Vaihingen 数据集上的精确率对比
Table 8 Precision comparison on the Vaihingen dataset

模型	精确率					
	路	建筑物	低矮植物	树	汽车	杂物
FCN	0.93	0.85	0.77	0.87	0.91	0.47
SegNet	0.91	0.76	0.84	0.90	0.80	0.55
U-Net	0.94	0.90	0.84	0.96	0.91	0.73
EaNet	0.93	0.94	0.81	0.94	0.87	0.77
MaNet	0.95	0.95	0.84	0.94	0.94	0.74
MF-FCNN	0.95	0.95	0.89	0.93	0.97	0.88

表9 Vaihingen 数据集上的召回率对比
Table 9 Recall comparison on the Vaihingen dataset

模型	召回率					
	路	建筑物	低矮植物	树	汽车	杂物
FCN	0.87	0.96	0.88	0.71	0.71	0.72
SegNet	0.89	0.96	0.89	0.75	0.70	0.65
U-Net	0.94	0.98	0.91	0.82	0.81	0.73
EaNet	0.95	0.97	0.86	0.86	0.80	0.79
MaNet	0.95	0.98	0.90	0.88	0.75	0.93
MF-FCNN	0.96	0.97	0.88	0.92	0.84	0.94

表10 Vaihingen 数据集上的平均 I_{ou} 和 O_A 对比
Table 10 Average I_{ou} and O_A comparison on the Vaihingen dataset

比较内容	平均 I_{ou}	$O_A/\%$
FCN	0.64	86.28
SegNet	0.68	85.27
U-Net	0.73	88.13
EaNet	0.78	90.79
MaNet	0.81	92.26
MF-FCNN	0.81	93.19

3.4 消融试验

为了验证 MF-FCNN 中 FLM 与 ASPP 的有效性,本研究在 Potsdam 数据集和 Vaihingen 数据集上

分别进行了消融试验,在 MF-FCNN 基础上,分别去除 FLM 与 ASPP。可视化试验结果如图 8、9 所示。其中,矩形框内为放大后的效果图。

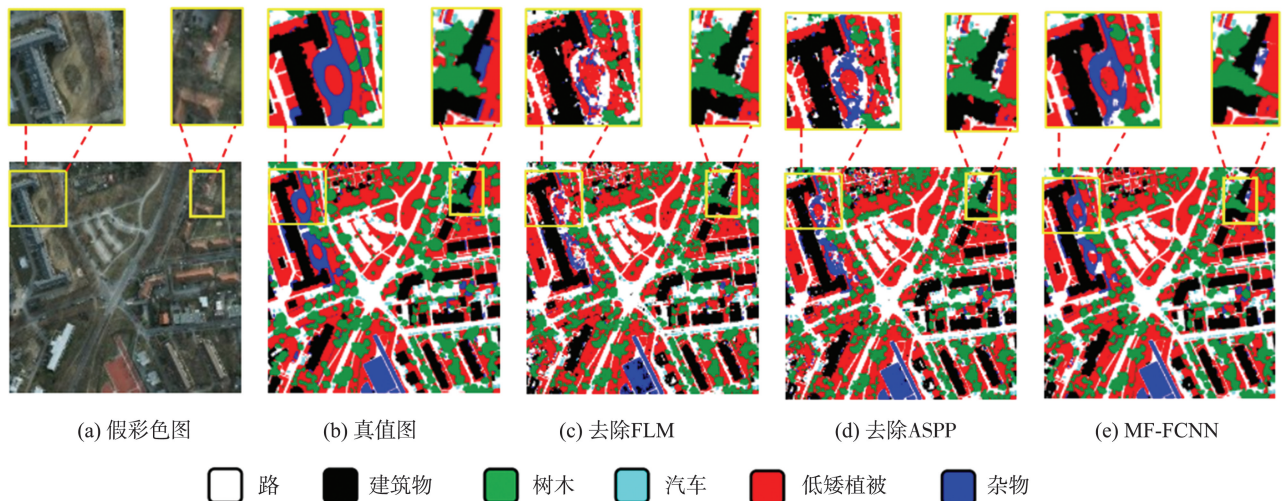


图8 Potsdam 数据集(全局)的消融试验结果
Fig.8 Ablation experiment results on the Potsdam dataset (global)

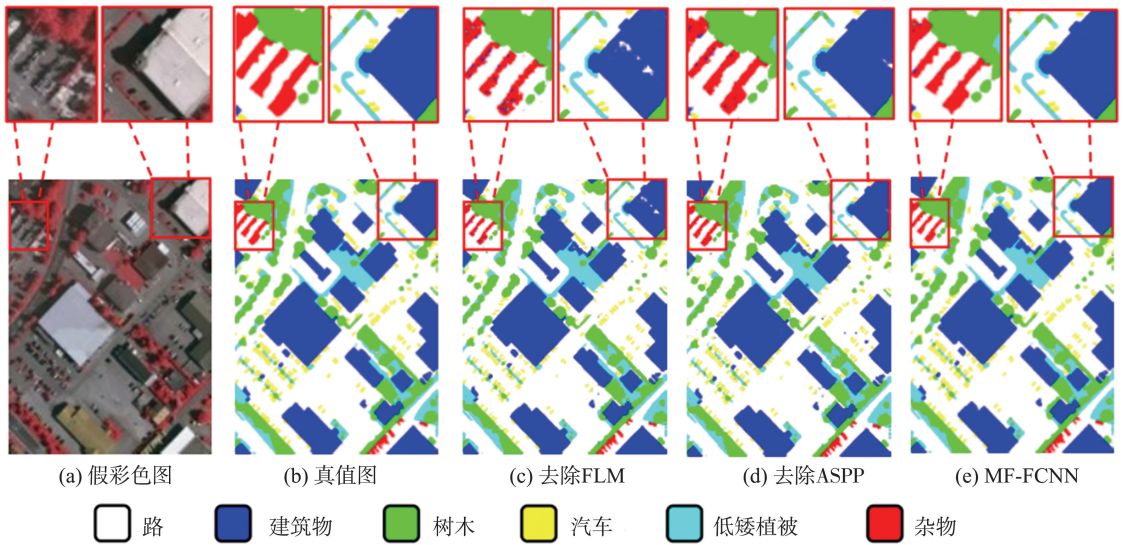


图9 Vaihingen数据集(全局)的消融试验结果
 Fig.9 Ablation experiment results on the Vaihingen dataset (global)

可视化试验结果显示,去除 FLM 后建筑物内部分割不完整且目标边缘分割粗糙现象明显,杂物类别的分割由于噪声的干扰存在较多点状物和错误分割情况。消融试验证明,FLM 能有效减少噪声对图像分割的影响,解决遥感图像中的不确定性问题,提高分割精度。另外,MF-FCNN 与未加 ASPP 的可视化结果图相比,分割效果也有一定程度提升,这也充分证明 ASPP 能够辅助提取多尺度信息,提升网络分割性能。

4 结语

本研究针对 HRI 中类内异质性、类间相似性的不确定性、网络浅层特征不足,获取上下文信息不完整等问题,提出一种基于多尺度特征的模糊卷积神经网络模型。该网络模型设计并使用 FLM 对不同尺度的特征进行模糊处理,学习高级语义,减少噪声干扰,解决 HRI 中固有的不确定性问题。聚合改进的 ASPP 处理多尺度目标并获取多尺度上下文信息。最后,在 Vaihingen 和 Potsdam 两个 HRI 数据集上进行了对比试验,试验结果证明,MF-FCNN 可以成功提取给定类别的每个对象,分类错误更少,边界更清晰。

参考文献:

[1] 侯月武,刘兆英,张婷,等. 基于改进的 DUNet 遥感图像道路提取[J]. 山东大学学报(工学版), 2022, 52(4): 29-37.
 HOU Yuewu, LIU Zhaoying, ZHANG Ting, et al. Road extraction from remote sensing images based on improved DUNet[J]. Journal of Shandong University (Engineering

Science), 2022, 52(4): 29-37.
 [2] KOUZIOKA G N, PERAKIS K. Decision support system based on artificial intelligence, GIS and remote sensing for sustainable public and judicial management[J]. European Journal of Sustainable Development, 2017, 6(3): 397-404.
 [3] SU T F, ZHANG S W. Local and global evaluation for remote sensing image segmentation[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 130(1): 256-276.
 [4] HUANG X, WEN D, LI J, et al. Multi-level monitoring of subtle urban changes for the megacities of China using high-resolution multi-view satellite imagery[J]. Remote Sensing of Environment, 2017, 196(1): 56-75.
 [5] ZHAO T Y, XU J D, CHEN R, et al. Remote sensing image segmentation based on the fuzzy deep convolutional neural network[J]. International Journal on Remote Sensing, 2021, 42(16): 6267-6286.
 [6] WANG Y, LIANG B, DING M, et al. Dense semantic labeling with atrous spatial pyramid pooling and decoder for high-resolution remote sensing imagery[J]. Remote Sensing, 2018, 11(1): 20-40.
 [7] 周力凯,江雨洋,冯亚春,等. 基于多尺度区域与类不确定性理论的局部阈值分割方法[J]. 计算机应用, 2020, 40(2): 66-72.
 ZHOU Likai, JIANG Yuyang, FENG Yachun, et al. Local threshold segmentation method based on multi-scale region and class uncertainty theory[J]. Journal of Computer Applications, 2020, 40(2): 66-72.
 [8] 曹春红,段鸿轩,曹玲,等. 基于多级特征级联的遥感图像实时语义分割[J]. 山东大学学报(工学版), 2021, 52(2): 19-25.
 CAO Chunhong, DUAN Hongxuan, CAO Ling, et al. Real-time semantic segmentation of high-resolution remote sensing image based on multi-level feature cascade[J].

- Journal of Shandong University (Engineering Science), 2021, 52(2): 19-25.
- [9] PANBOONYUEN T, JITKAJORNWANICH K, LA-WAWIROHWONG S, et al. Semantic segmentation on remotely sensed images using an enhanced global convolutional network with channel attention and domain specific transfer learning [J]. Remote Sensing, 2019, 11(1): 83-105.
- [10] KONSTANTINOVA T, WIEGART L, RAKITIN M, et al. Noise reduction in X-ray photon correlation spectroscopy with convolutional neural networks encoder-decoder models [J]. Scientific Reports, 2021, 11(1): 1-12.
- [11] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [12] 张小娟, 汪西莉. 完全残差连接与多尺度特征融合遥感图像分割 [J]. 遥感学报, 2020, 24(9): 1120-1133. ZHANG Xiaojuan, WANG Xili. Image segmentation models of remote sensing using full residual connection and multiscale feature fusion [J]. Journal of Remote Sensing, 2020, 24(9): 1120-1133.
- [13] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Proceedings of the European Conference on Computer Vision. Munich, Germany: Springer, 2018: 801-818.
- [14] 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.
- [15] DU S J, DU S H, LIU B, et al. Incorporating DeepLabv3+ and object-based image analysis for semantic segmentation of very high resolution remote sensing images [J]. International Journal of Digital Earth, 2021, 14(3): 357-378.
- [16] FENG Y C, DIAO W H, SUN X, et al. NPALoss: neighboring pixel affinity loss for semantic segmentation in high-resolution aerial imagery [J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, 20(2): 475-482.
- [17] TANG M, DJELOUAH A, PERAZZI F, et al. Normalized cut loss for weakly-supervised CNN segmentation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE, 2018: 1818-1827.
- [18] ZHU L Y, JI D Y, ZHU S P, et al. Learning statistical texture for semantic segmentation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE, 2021: 12537-12546.
- [19] TABRIZI P R, MANSOOR A, CERROLAZA J J, et al. Automatic kidney segmentation in 3D pediatric ultrasound images using deep neural networks and weighted fuzzy active shape model [C]// Proceedings of 2018 IEEE 15th International Symposium on Biomedical Imaging IEEE. Piscataway, USA: IEEE, 2018: 1170-1173.
- [20] HASAN P. Multi-task semantic segmentation of CT images for COVID-19 infections using DeepLabV3+ based on dilated residual network [J]. Physical and Engineering Sciences in Medicine, 2022, 12(4): 1-13.
- [21] WANG Q, ZHANG X, CHEN G, et al. Change detection based on faster R-CNN for high-resolution remote sensing images [J]. Remote Sensing Letters, 2018, 9(10): 923-932.
- [22] DENG G H, WU Z C, WANG C J, et al. CCANet: class-constraint coarse-to-fine attentional deep network for subdecimeter aerial image semantic segmentation [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 2(99): 1-20.
- [23] LI R, ZHENG S, DUAN C. Multistage attention ResUNet for semantic segmentation of fine-resolution remote sensing images [J]. IEEE Geoscience and Remote Sensing Letters, 2022, 40(19): 1-5.
- [24] SU Y, CHENG J, BAI H W, et al. Semantic segmentation of very-high-resolution remote sensing images via deep multi-feature learning [J]. Remote Sensing, 2022, 14(3): 533-558.
- [25] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [J]. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015, 17(12): 234-241.
- [26] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(12): 2481-2495.
- [27] ZHENG X, HUAN L, XIA G S, et al. Parsing very high resolution urban scene images by learning deep ConvNets with edge aware loss [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 170(1): 15-28.
- [28] LI R, ZHENG S, DUAN C, et al. Multiattention network for semantic segmentation of fine-resolution remote sensing images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60(1): 1-3.
- [29] ROTTENSTEINER F, SOHN G, JUNG J, et al. The ISPRS benchmark on urban object classification and 3D building reconstruction [J]. Remote Sensing and Spatial Information Sciences, 2012, 2(5): 293-298.