引文格式：张京奎,产桂洋,陈曙光,等. 结合超分辨率重建技术的遥感语义分割方法[J]. 地理空间信息, 2025, 23(1): 29-33. doi:10. 3969/j.issn. 1672-4623.2025.01.007

结合超分辨率重建技术的遥感语义分割方法

张京奎[1,2]，产桂洋[3*]，陈曙光[4]，刘　臻[1,2]

（1. 沛县自然资源和规划局，江苏 徐州 221699；2. 自然资源部国土空间大数据工程技术创新中心 应用技术研发分中心，江苏 徐州 221699；3. 河海大学 计算机与软件学院，江苏 南京 210024；4. 特力惠信息科技股份有限公司，福建 福州 350706）

摘　要：针对遥感数据空间分辨率低、制作高精度样本标签难度大、样本数量不足等问题，提出了一种基于迁移学习思想结合超分辨率重建技术的遥感图像语义分割方法。将该方法应用于沛县建筑物提取中，结果表明提取的平均交并比可达72.4%，比不做任何处理提升了22.3%，说明该方法具有较高的分割精度、效率和可用性，为遥感图像目标提取提供了一种新的思路和解决方案，具有实际应用价值和研究意义。

关键词：遥感影像；迁移学习；语义分割；超分辨率重建

中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2025）01-0029-05

Semantic Segmentation Method for Remote Sensing Images Using Super-resolution Reconstruction Technology

ZHANG Jingkui[1,2], CHAN Guiyang[3], CHEN Shuguang[4], LIU Zhen[1,2]

(1. Pei County Natural Resources and Planning Bureau, Xuzhou 221699, China;
2. The Application Technology R&D Sub-center, National Land Spatial Data Engineering Technology Innovation Center, Ministry of Natural Resources, Xuzhou 221699, China;
3. College of Computer Science and Software Engineering, Hohai University, Nanjing 210024, China;
4. Techwire Information Technology Co., Ltd., Fuzhou 350706, China)

Abstract: In this paper, we proposed a remote sensing image semantic segmentation method based on transfer learning, combining super-resolution reconstruction techniques to address the problems of low spatial resolution, difficulty in creating high-accuracy sample labels, and insufficient sample quantity. We applied this method to the building extraction in Pei County. The result show that the average IoT ratio of extracted buildings reaches 72.4%, which is a 22.3% improvement compared to not doing any processing, which proves that the method has high segmentation accuracy, efficiency and usability, providing a new approach and solution for remote sensing image object extraction, with practical application value and research significance.

Key words: remote sensing image, transfer learning, semantic segmentation, super-resolution reconstruction

随着遥感技术的不断发展和应用范围的扩大，遥感数据的获取和处理已成为各领域的研究热点。遥感图像分割[1]是遥感数据处理中一个重要环节，其核心是将遥感图像中的每个像素归类到不同的目标区域中。为提高遥感图像分割的准确性和鲁棒性，学者们提出了基于深度学习的分割方法，可以自动学习特征，适应各种遥感图像的分割任务。然而，由于遥感图像中涉及大量地物、地貌、气象等复杂信息，因此制作高质量的遥感数据集十分困难。基于此，学者们开始探索利用迁移学习方法来提高遥感图像处理的效率。迁移学习[2]是一种将已学到的知识迁移到新任务中的机器学习方法。在遥感图像处理中，利用迁移学习方法可将已经训练好的模型或特征提取器应用于新的遥感图像处理任务中，从而节省大量的数据和计算资源。近年来针对遥感图像的迁移学习已成为研究的重要领域，Dong R C 等[3]将 Sig-NMS-based Faster R-CNN 与迁移学习相结合，以解决缺乏已建立的包含标记的小目标数据集的问题；Zhao H 等[4]将 CNN 迁移到遥感场景分类中，以修正目标与源数据集之间的输入特征差异；Cui B G 等[5]尝试利用迁移学习和新型深度卷积神经网络进行语义分割；Liu Y[6]等利用迁移学习进行语义分割，可从相关领域的数据中提取有用信息并将其转移到目标领域，以解决跨域学习的问题。尽管上述方法能大幅减少遥感数据集制作的工作量，

第一作者简介：张京奎（1970—），本科，研究方向为国土信息化、国土大数据、土地管理，E-mail：763053655@qq.com。
通信作者：产桂洋（2000—），硕士研究生，主要研究方向为深度学习、人工智能、遥感图像处理，E-mail：2469362360@qq.com。

并得到了不错的预测效果，但仍存在一定局限性：①当公共数据集与目标研究区的地理样貌存在巨大差异时，训练完成的模型对研究区内测试数据的预测效果可能达不到预期；②现有公共数据集空间分辨率较高，而对于空间分辨率较低的测试数据，由于特征尺度不一致，直接使用迁移学习的方法预测的精度并不高。

本文针对遥感数据空间分辨率低、制作高精度样本标签难度大、样本数量不足等问题，提出了一种基于迁移学习思想结合超分辨率重建技术的语义分割方法。通过预训练模型和数据增强，可有效增加样本的多样性和数量，进而提高语义分割模型的准确性和鲁棒性；通过超分辨率重建技术将低分辨率图像转换为高分辨率图像，可进一步提高分割精度。本文将该方法应用于沛县建筑物提取的实际任务中，结果表明该方法具有较高的分割精度、效率和可用性。

1 研究数据与方法

1.1 研究区概况

本文选取沛县作为研究区。沛县位于中国江苏省徐州市的西北部、苏鲁两省交界处，毗邻山东省鱼台县，东临微山湖、西接丰县、南邻铜山区，是华北平原东南缘和淮海经济区西北部的一部分；地理坐标为34°43′~34°57′N、116°44′~117°04′E；总面积为1 448 km^2，下辖18个乡镇、1个街道办事处和1个开发区，人口约为97万人；地势东北高、西南低，地形以平原为主，境内沙河、白河、北淝河、南淝河等河流纵横交错。

1.2 数据来源

本文采用的公共数据集是中国地质大学（武汉）方芳团队以高分辨率遥感影像为数据源[7]，采用人工标注与交互式标注相结合的方式构建的中国典型城市建筑物实例数据集。数据集包含7 260个影像区域样本，共63 886栋建筑物，分布在北京、上海、深圳和武汉4个城市；由MS COCO 2017格式的标注文件及其相应的建筑物掩膜二值图构成。结合沛县的现有数据情况，本文采用沛县空间分辨率为0.1 m和1 m的遥感卫星影像数据，其中空间分辨率为0.1 m的遥感影像用于训练超分辨率重建模型，但由于公有数据集与现有遥感影像在分辨率上仍存在差异，因此对其进行下采样制作成接近公共数据集分辨率的数据；空间分辨率为1 m的遥感影像用于数据增强实验和推理测试。测试数据集为利用沛县空间分辨率为1 m的遥感影像数据自行人工制作。

1.3 技术方案

本文提出了一种结合超分辨率重建的语义分割方法，用于分割和预测高质量图像，并利用风格迁移技术增强样本。该方法整体采用迁移学习的思想，由数据增强、语义分割和超分辨率重建3个模块组成。数据增强模块利用风格迁移模型对现有高分辨率公共数据集进行处理，生成类似于目标测试区域的数据样本，其目的是使模型适应并学习研究区所需的特定风格特征。语义分割模块首先利用公共数据集进行预训练，获取初始的模型参数和特征表示；再利用风格迁移后的数据样本进行微调训练，得到适合于预测研究区的语义分割模型。超分辨率重建模块利用训练好的超分辨率重建模型将低分辨率的研究区图像转换为高分辨率图像，可为后续的推理预测提供更高质量的输入图像。

1.4 关键技术

1）语义分割。语义分割[8]是一种计算机视觉技术，旨在将图像中的每个像素分配到相应的语义类别中。与传统图像分类和目标检测技术不同，语义分割不仅能识别出图像中的物体，还能精确地分割每个物体的轮廓和边界。语义分割依赖于深度学习模型，特别是卷积神经网络（CNN）[9-11]。CNN模型能自动地从原始图像中提取特征，并将这些特征用于像素级别的分类。遥感图像语义分割可从遥感图像中精确提取每个像素的语义信息，从而更好地理解地表特征和变化。InternImage[12]是一种基于CNN的模型，旨在突破传统CNN模型在扩大规模和处理长距离依赖方面的限制。与最近的大核卷积和Transformer模型不同，InternImage采用可变性卷积作为核心算子，以实现大尺度有效感受野和自适应空间聚合；采用DCNv3算子，引入了共享投射权重、多组机制和采样点调制等改进；通过设计动态稀疏卷积算子，实现全局注意力效果，同时降低计算和存储资源的消耗，以实现高效训练；基础模块采用了GELU、层归一化和前馈网络等先进组件，使得其在各种视觉任务中的表现更加高效。

2）超分辨率重建。训练好语义分割模型后，需要对测试区域的遥感影像进行预测；但测试区域的遥感影像与现有训练数据之间的空间分辨率存在差异，将影响模型的预测效果。为解决该问题，本文采用超分辨率重建技术[13]将低分辨率图像扩大到高分辨率。超分辨率重建技术可通过插值、CNN等方法，将低分辨率图像转换到与现有训练数据相同的特征尺度，提

高匹配和对齐的效果。此外，更高的分辨率有利于语义分割模型提取更准确、更细节的特征，从而提高模型的准确性和泛化能力。Real-ESRGAN[14]是一种基于生成对抗网络（GAN）的超分辨率图像重建方法，包括生成器和判别器两个关键网络。生成器通过多个卷积层、残差块和上采样层将低分辨率图像转换为高分辨率图像；判别器则是一个二进制分类器，用于区分生成的高分辨率图像和真实图像。训练时使用3种损失函数对抗损失驱使生成器生成逼真图像，内容损失确保生成器输入图像的视觉内容，特征感知损失有助于捕捉更高层次的语义信息，以提高图像的感知质量。

2　实验与结果分析

2.1　实验方法

根据本文的方法思路，实验的具体步骤为：①预训练，利用公共数据集对InternImage模型进行预训练，以获取初始的模型参数和特征表示；②数据增强，通过CycleGAN[14]模型对公共数据和研究区数据进行对抗训练，生成研究区域特定风格的公共数据；③微调语义分割模型，利用经过风格迁移后的数据进行监督学习，对初步训练后的InternImage模型进行微调训练，旨在实现更精确的像素级别分割，并学习和适应研究区域的风格特征；④训练超分辨率重建模型Real-ESRGAN，利用研究区高分辨率数据训练超分辨率重建模型，该模型具有将低分辨率图像转换为高分辨率图像的能力，从而提升图像的细节和清晰度；⑤推理预测，先将低分辨率研究区图像输入已训练好的超分辨率重建模型中，以获得高分辨率图像，再将高分辨率图像输入到训练好的语义分割模型中进行预测，以获得最终的结果，结合了超分辨率重建和语义分割的能力，实现了对研究区图像的高质量分割和预测。

2.2　实验过程

1）预训练。本文采用中国典型城市建筑物实例数据集对InternImage模型进行预训练。模型基于MMCV框架实现，采用UperNet_r50架构，以ResNet-50为特征提取器，采用U-Net架构进行特征融合，同时利用其特有的动态卷积提升模型的特征提取和语义理解能力。训练前对公有数据集进行归一化处理、数据预处理和增强，并以7∶3的比例划分训练集和验证集。模型采用AdamW优化器进行训练，学习率设置为0.000 06，权重衰减系数为0.05，采用多项式衰减策略调度学习率，并采用线性预热。将Batch Size设置为8，模型进行20 000次迭代训练，每1 000轮进行一次模型评估，评估指标为平均交并比，最终选择验证集上表现最好的模型作为结果。

2）数据增强。考虑到不同地理位置上遥感影像特征风格的差异性，为进一步扩大样本数量、增强语义分割模型对目标域的域适应能力，本文将其设置为源域数据，利用CycleGAN模型将其中的样本风格迁移到目标域（沛县）风格。实验过程中共训练100轮，Batch Size设置为1，学习率设置为0.000 2，采用学习率衰减策略，从70轮开始直至训练结束，学习率逐渐线性衰减到0。训练完毕后，当总损失趋近于收敛时，冻结CycleGAN中生成器的权重，以公有数据集中的样本作为输入，将其全部转化为沛县风格的样本。

3）微调。预训练完成后，采用数据增强后的数据集对语义分割模型进行微调，以适应研究区的风格和特点。由于预训练模型已具有较强的泛化能力，在新数据集上表现良好，因此微调过程中保留预训练模型中的大部分参数，仅微调最后几层，以适应新数据集中的特定特征和语义类别；同时调整超参数以优化模型在新数据集上的性能。微调完成后，利用训练好的语义分割模型进行下一步超分辨率重建模型的训练和推理测试。

4）超分辨率重建。超分辨率重建采用研究区空间分辨率为0.1 m的遥感影像对Real-ESRGAN进行训练。训练过程中先利用Real-ESRGAN中的降级模型生成低质量图像，然后在已有预训练模型权重的基础上进行微调训练。生成器采用RRDBNet，判别器采用带有跳跃连接的U-Net，生成器和判别器的优化器均采用Adam，学习率设置为0.000 1，学习率调整器采用多步长学习率调整器。L1损失、感知损失和GAN损失的权重分别设置为1、1和0.1。微调训练迭代次数设置为80 000次。

5）推理预测。在完成超分辨率重建模型和语义分割模型训练后，可对研究区遥感影像进行推理测试。首先将低分辨率的研究区图像输入训练好的超分辨率重建模型中进行处理，得到高分辨率图像；再将高分辨率图像输入训练好的语义分割模型中进行预测，得到每个像素点属于不同语义类别的概率分布。对于每个像素点，选择概率最大的语义类别作为其预测结果，从而得到整张高分辨率图像的语义分割结果。

2.3　实验结果分析

2.3.1　评价指标

1）交并比（IoU）。在语义分割任务中，IoU是一种

常用的评价指标，用于衡量模型预测的分割结果与真实分割结果之间的重叠程度。假设模型预测的分割结果为 P，真实分割结果为 T，则有：

$$\mathrm{IoU}=\left|\frac{P\cap T}{P\cup T}\right| \tag{1}$$

式中，$P\cap T$、$P\cup T$ 分别为预测的分割结果与真实分割结果的交集区域和并集区域。

平均交并比（MIoU）是各类别IoU的平均值。IoU和MIoU的取值范围为0~1，值越高表示模型的性能越好。

2）F1-score在语义分割任务中，F1-score是像素级别的评估指标，用于衡量模型对每个像素预测结果的准确性。F1-score可被视为像素级别的精确度（Precision）和召回率（Recall）的调和平均数。对于每个像素，首先计算其预测类别与真实类别之间的交集和并集，再计算它们的比率，即像素级别的Precision和Recall，最后计算所有像素的F1-score，即

$$F1=\frac{2\times\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}} \tag{2}$$

F1-score的取值通常为0~1，越接近1表示模型的性能越好。

2.3.2 对比方法设置

为了更全面地评估本文方法的效果，将实验分为5组：真实结果、直接训练公共数据集再预测的结果、只对公共数据集进行数据增强再预测的结果、只对公共数据集进行超分辨率重建再预测的结果以及本文方法预测结果。首先采用人工标注的方式对原始研究区内的图像进行标注，得到真实的分割结果；再以手动标注的样本标签为测试集对其余4组实验预测结果进行测试，得到评价指标和预测效果图。

2.3.3 实验结果与分析

在数据增强实验中（图1），风格迁移后的数据在图像风格特征方面向沛县的图像风格接近，为后续的语义分割模型训练提供了良好的数据基础；在超分辨率重建实验中，对沛县的图像数据进行了超分辨率重建，分辨率有了显著提升，一些细节特征也得到了改善，重建后的影像数据在特征尺度上也更接近于公共数据。

通过测试得到的实验结果见表1，可以看出，采用CycleGAN和Real-ESRGAN技术可提高建筑物和背景的像素分类性能，从而提高MIoU和F1-score。具体来说，利用CycleGAN技术可将建筑物、背景的预测结果分别提高8%和7.7%；利用Real-ESRGAN技术可将建筑物、背景的预测结果分别提高16.4%和16.3%；而

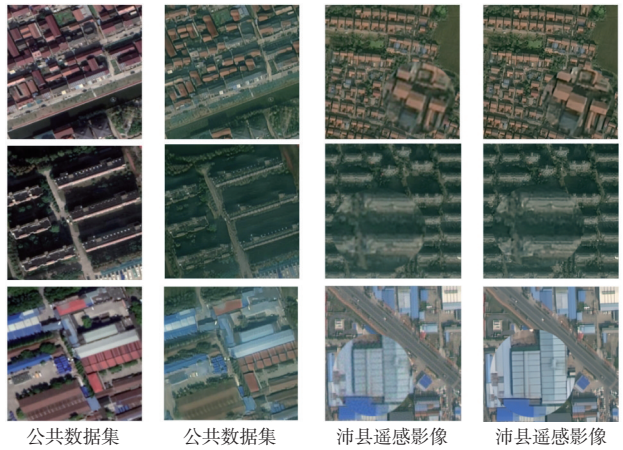


图1 数据增强和超分辨率重建前后对比图

综合利用两种技术可进一步提高预测的像素分类性能，将建筑物、背景的预测结果分别提高15.2%和16.6%。综上所述，CycleGAN和Real-ESRGAN技术可有效提高语义分割模型的性能，在提高建筑物和背景的像素分类性能方面发挥了重要作用，从而提高了MIoU和F1-score指标。

表1 沛县数据集消融实验/%

方法	建筑物	背景	MIoU	F1-score
NoAdapt	33.7	66.4	50.1	48.6
CycleGAN	41.5	74.1	57.8	34.7
Real-ESRGAN	50.1	82.7	66.4	62.3
CycleGAN+ Real-ESRGAN	56.3	88.5	72.4	71.1

沛县遥感图像的预测结果见图2，可以看出，数据增强后的预测结果比未增强的结果在细节上有所提升，在公共数据集与沛县实际风格不一致的图像上尤为明显；超分辨率重建后的预测结果性能提升非常明显，如图2第二排和第四排图像，一些建筑在超分辨率重建后能被模型正确识别，这是由于分辨率的提升使语义分割模型能更好地提取特征，从而得到更好的结果。

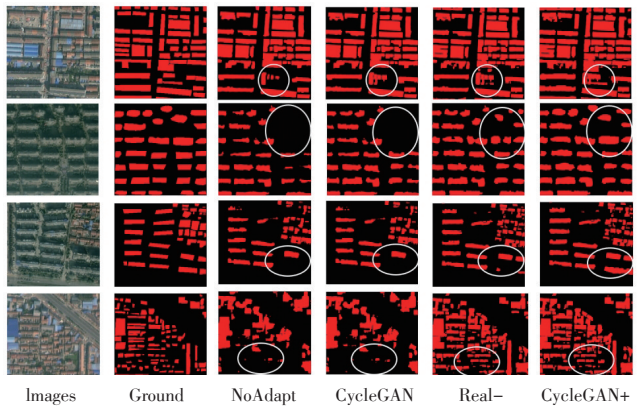


图2 沛县遥感数据预测结果图

3 结 语

本文提出了一种新的结合超分辨率重建技术的遥感图像语义分割方法。该方法在沛县建筑物提取任务中得到了验证。结果表明，该方法具有较高的分割精度、效率和可用性。该方法为遥感图像语义分割问题提供了新的解决方案，克服了遥感数据空间分辨率低、制作高精度样本标签难度大和样本数量不足等问题，具有实用价值和研究意义。其优越性在于将语义分割和超分辨率重建技术相结合，能够提高图像质量和分割精度，并保持较高的效率和可用性。未来将进一步完善和优化该方法，以提高分割精度和鲁棒性，并将其应用于更广泛的领域和任务中。本文提出的方法为遥感图像语义分割问题提供了一种新思路，有望在实际应用中发挥重要作用。

参考文献

[1] 闵蕾,高昆,李维,等. 光学遥感图像分割技术综述[J]. 航天返回与遥感,2020,41(6):1-13

[2] 赵凯琳,靳小龙,王元卓. 小样本学习研究综述[J]. 软件学报,2021,32(2):349-369

[3] Dong R C, Xu D Z, Zhao J, et al. Sig-NMS-based Faster R-CNN Combining Transfer Learning for Small Target Detection in VHR Optical Remote Sensing Imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11):8 534-8 545

[4] Zhao H, Liu F, Zhang H, et al. Convolutional Neural Network Based Heterogeneous Transfer Learning for Remote Sensing Scene Classification[J]. International Journal of Remote Sensing,2019(16):156-162

[5] Cui B G, Chen X, Lu Y. Semantic Segmentation of Remote Sensing Images Using Transfer Learning and Deep Convolutional Neural Network with Dense Connection[J]. IEEE Access,2020,8(6):116 744-116 755

[6] Liu Y, Kong Y, Zhang B, et al. A Novel Deep Transfer Learning Method for Airborne Remote Sensing Semantic Segmentation Based on Fully Convolutional Network [C]. 4th International Conference on Imaging, Signal Processing and Communications (ICISPC),2020

[7] Fang F, Wu K, Zheng D. A Dataset of Building Instances of Typical Cities in China(V2)[DB]. 2021-03-25

[8] 邝辉宇,吴俊君. 基于深度学习的图像语义分割技术研究综述[J]. 计算机工程与应用,2019,55(19):12-21

[9] 刘蝶. 基于卷积神经网络的航空影像城市建筑物分割[J]. 地理空间信息,2020,18(1):51-53

[10] 陆妍如,毛辉辉,贺琰,等. 基于Deeplab v3+的高分辨率遥感影像地物分类研究[J]. 地理空间信息,2022,20(6):1-6

[11] 石婷婷,戴腾,厉芳婷,等. 基于深度学习的耕地“非农化”遥感监测初探[J]. 地理空间信息,2022,20(3):34-37

[12] Wang W, Dai J, Chen Z, et al. Internimage: Exploring Large-scale Vision Foundation Models with Deformable Convolutions[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2023:14 408-14 419

[13] 张良培,黄昕. 遥感影像信息处理技术的研究进展(英文)[J]. 遥感学报,2009,13(4):559-569

[14] Zhu J Y, Park T, Isola P, et al. Unpaired Image-to-Image Translation Using Cycle-consistent Adversarial Networks[C]. IEEE International Conference on Computer Vision (ICCV),2017

（上接第20页）

[7] 张宏,贺也平,石志国. 一个支持空间上下文的访问控制形式模型[J]. 中国科学(E辑:信息科学),2007,37(2):254-271

[8] Damiani M L, Bertino E, Catania B, et al. GEO-RBAC: GEO-RBAC: A Spatially Aware RBAC[J]. ACM Transactions on Information and System Security,2007,10(1):2

[9] 张颖君,冯登国. 基于尺度的时空RBAC模型[J]. 计算机研究与发展,2010,47(7):1 252-1 260

[10] 祁琼,赖云,周宝玉. 国土调查移动数字管理系统关键技术研究[J]. 地理空间信息,2023,21(2):72-76

[11] 徐杨. 空间数据访问控制关键技术研究[D]. 郑州:信息工程大学,2012

[12] 毛健. 地理空间数据访问控制模型研究[D]. 南京:南京师范大学,2015

[13] Zhang Y, Zhang Y, Chen K. A Map-layer-based Access Control Model[C]. Information Security Applications: 12th International Workshop,Jeju Island,2011

[14] Covington M, Long W, Srinivasan S, et al. Securing Context-aware Applications Using Environment Roles[C]// Proceedings of the Sixth ACM Symposium on Access Control Models and Technologies,2001

[15] 尹峰,李俭康,刘中秋. 湖北省级国土调查数据库优化与关键技术研究[J]. 地理空间信息,2022,20(5):18-21

[16] 蔡丽娟,陈伟鹤. 带有时空特性的角色访问控制模型[J]. 软件导刊,2017,16(8):15-18

[17] Bertolissi C, Fernandez M. Time and Location Based Services with Access Control[C]. IEEE New Technologies, Mobility and Security,2008

[18] 安小明,王小明,王巧玲. 具有时空约束的角色访问控制模型[J]. 计算机工程与应用,2010,46(7):89-92