引文格式：李昊燔,黎娟,孔令寅. 交叉验证的半监督遥感图像分割方法[J]. 地理空间信息, 2026, 24(4):126-129. doi:10. 3969/j.issn. 1672-4623.2026.04.026

交叉验证的半监督遥感图像分割方法

李昊燔，黎　娟，孔令寅
（西安航空职业技术学院 通用航空学院，陕西 西安 710089）

摘　要：现有的遥感图像处理方法依赖于有监督训练，模型精度严重依赖于标签数据。基于现有的深度学习编码-解码语义分割架构，提出一种双模型交叉验证的半监督遥感图像语义分割方法。首先利用少量标签图像对两组模型进行有监督训练迭代模型参数，其次大量的无标签数据分别通过两组模型得到对应伪标签，最后利用伪标签进行模型间交叉监督训练。实验结果表明，该方法利用了大量的无标签数据用于深度学习模型训练，有效降低了模型训练所需的标签成本。

关键词：遥感图像；语义分割；半监督；交叉验证
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2026）04-0126-04

Semi-supervised Remote Sensing Image Segmentation Method Based on Cross-validation

LI Haofan, LI Juan, KONG Lingyin
(School of General Aviation Institute, Xi'an Aeronautical Polytechnic Institute, Xi'an 710089, China)

Abstract: Most of the remote sensing image processing methods rely on supervised training, and the model accuracy depends on the labeled data. Based on the deep learning encoding-decoding semantic segmentation architecture, we proposed a semi-supervised remote sensing image semantic segmentation method with two-model cross-validation. Firstly, we used a small number of labeled images to iterate the model parameters for supervised training. And then, we let the unlabeled data pass through the two models to get the pseudo-labels. Finally, we used the pseudo-labels to perform cross-supervised training. Experimental result shows that the method uses the unlabeled data to train deep learning model, effectively reducing the training cost.

Key words: remote sensing image, semantic segmentation, semi-supervised, cross-validation

近年来，随着半导体技术的不断进步[1-10]，遥感成像的便捷性得到了显著提升[11]，从而产生了大量的无标注遥感数据。然而，这些无标注数据的潜力尚未得到充分发挥。如何有效利用这些无标签数据来提升图像分割模型的性能，成为了当前深度学习领域的一个重要研究方向[12]。针对这一挑战，研究者们正在探索半监督学习等新兴方法，以在减少对标注数据依赖的同时，挖掘无标签数据中蕴含的信息，进而优化遥感图像的解析能力[13-14]。这些研究不仅有助于降低数据标注成本，还能够提高模型在复杂环境下的适应性和准确性。

1　本文模型及算法概述

1.1　UNet语义分割网络

UNet网络最早在医学图像处理领域得到了广泛应用，这得益于其对称的编码-解码结构以及通过跳跃连接实现的编码层与对应解码层的有效对接。这种设计不仅确保了模型在解码过程中能够充分利用上下文信息，而且显著提升了语义分割的预测精度。UNet网络的对称结构在遥感图像分割领域也展现出了卓越的表现。特别是其跳跃连接机制，极大地改善了对遥感图像中细微特征的识别能力，从而提高了特征不明显类别的分割效果。

1.2　交叉验证的半监督方法

图像一致性原则[14]指出，在相同输入图像经过不同处理程度后，其在任何像素位置的预测结果应保持一致。这一原则为半监督学习方法的设计提供了重要理论依据。基于图像一致性假设，本文提出了一种改进的交叉验证半监督遥感图像分割方法。该方法利用两组结构相同但预处理方式不同的UNet分割模型进行训练，并在一个包含有限标签数据和大量无标签数据的数据集中应用。

项目来源：全国教育科学规划重点课题资助项目（DJA200310）；陕西省教育科学规划课题（SGH23Y3003）；西安航空职业技术学院校级课题（23XHZK-24）；西安航空职业技术学院2021年度科技创新团队资助项目（KJTD21-001）。
第一作者简介：李昊燔（1993—），硕士，讲师，研究方向为无人机航空摄影测量、遥感图像处理，E-mai：810677934@qq.com。

本研究使用交叉熵（Cross Entropy）作为损失函数进行损失计算。像素级的交叉熵损失[11]作为语义分割最经典的损失函数，能够单独考察每个像素，将每个像素类别的预测值与 标签进行比较。交叉熵由公式（1）定义。

$$CE(y_i, y_i^{'}) = -\sum_i ((y_i \log(y_i^{'}) + (1-y_i)\log(1-y_i^{'})) \quad (1)$$

对于包含 N 个大小为 $W \times H$ 的标签数据集 D^L，两组模型的监督损失，计算公式如式（2）、（3）：

$$L^{\mathrm{UNet1}} = \frac{1}{N} \sum_{x \in D^L} \frac{1}{W \times H} \sum_{i=1}^{W \times H} L_{CE}(P_i^{\mathrm{UNet1}}, Y_i^*) \quad (2)$$

$$L^{\mathrm{UNet2}} = \frac{1}{N} \sum_{x \in D^L} \frac{1}{W \times H} \sum_{i=1}^{W \times H} L_{CE}(P_i^{\mathrm{UNet2}}, Y_i^*) \quad (3)$$

式中，i 为输出图像的任一像素点，代表预测结果 P_i^{UNet1}；P_i^{UNet2} 为在像素点 i 在两组模型的预测值；L_{CE} 为公式（1）中提到的交叉熵损失函数；Y_i^* 为标签中像素点 i 的真实类别。

该计算方式保证了模型能够在有标签数据的指导下逐层优化其参数。最终，通过反向传播算法，模型的参数在有监督学习条件下被逐层修正，以实现更精确的图像分割。

本文中的监督学习训练主要有 2 个目标：一是引导模型在初始化阶段的训练，二是在后续训练过程中修正模型参数。为了实现这一目标，本文采用了一种交叉监督训练方法，该方法通过两组模型之间的交互学习，促使模型主动探索和利用无标签数据中的潜在信息。具体来说，这种交叉监督训练方法的流程如图 1 所示。

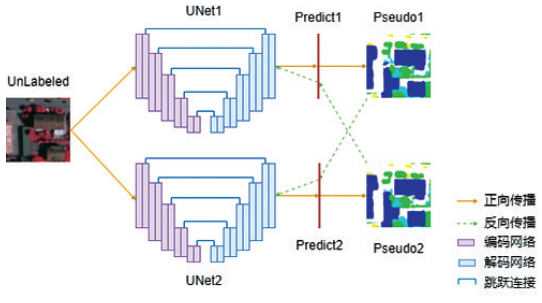


图 1　无标签数据训练过程

对于无标签数据，本文首先将输入图像分别输入两组训练好的 UNet 模型（UNet1 和 UNet2）进行预测，从而获得两组预测结果 Predict1 和 Predict2。根据这些预测结果，可以生成两组伪标签 Pseudo1 和 Pseudo2，这些伪标签是通过预测结果与具体像素值的映射关系推导出的。

基于图像一致性假设原则，即伪标签 Pseudo1 和 Pseudo2 应接近于图像的真实标签。本文利用这一假设，采用了交叉监督训练策略。具体方法是将伪标签 Pseudo1 和 Pseudo2 分别作为各自模型的标签数据，并计算模型预测结果与伪标签之间的交叉监督损失。

对于包含 M 个大小为 $W \times H$ 的无标签数据集 D^U，两组模型的交叉监督损失，计算公式如式（4），（5）：

$$L^{\mathrm{UNet2-UNte1}} = \frac{1}{M} \sum_{x \in D^U} \frac{1}{W \times H} \sum_{i=1}^{W \times H} L_{CE}(P_i^{\mathrm{UNet1}}, P_i^{\mathrm{UNet2}}) \quad (4)$$

$$L^{\mathrm{UNet1-UNte2}} = \frac{1}{M} \sum_{x \in D^U} \frac{1}{W \times H} \sum_{i=1}^{W \times H} L_{CE}(P_i^{\mathrm{UNet2}}, P_i^{\mathrm{UNet1}}) \quad (5)$$

通过反向传播算法，对两组模型的参数进行逐层修正，以最小化交叉监督损失，从而实现模型间的交叉监督训练。这种训练方式不仅增强了模型对无标签数据的学习能力，还提高了对伪标签质量的依赖，从而进一步提升了模型的性能和对遥感图像的分割精度。

2　实验与分析

2.1　实验环境

本实验使用的硬件环境包括 CPU：Intel Core i5-12490F 和 GPU：NVIDIA GeForce RTX 2080 Ti。软件环境方面，操作系统为 Ubuntu 18.04，深度学习框架为 PyTorch 1.9.0。这些配置提供了强大的计算能力和灵活的开发环境，以支持深度学习模型的训练和测试。

2.2　数据集及预处理

本研究选择国际摄影测量与遥感学会（ISPRS）提供的 Vaihingen 数据集作为验证对象，旨在评估所提出的交叉验证半监督遥感图像分割方法的有效性[12]。其次本研究在预训练阶段使用了同为国际摄影测量与遥感学会的 Potsdam 数据集为两组模型（UNet1 和 UNet2）生成了预训练权重，并且使用了不同的数据集进行权重生成，本研究将 Potsdam 数据集选取了自然要素居多的植被、水体数据集 200 张（256 像素×256 像素）以及人文要素居多的建筑、车体数据集 200 张（256 像素×256 像素），并使用两组数据集为两组模型分别生成预训练权重，这样可以保持两组模型在初始化阶段对于输入图像的不同关注度，从而为交叉监督提供更好的优化效果。

在实验环境的设置方面，使用高性能 GPU（如 NVIDIA GeForce RTX 2080 Ti），配合至少 16 GB 的 RAM，以确保在处理大规模图像数据时，模型能够高效运行。实验系统基于 Windows 操作平台，并采用了深度学习框架 TensorFlow 以及 PyTorch 进行模型训练，

具体实验环境信息参考2.1所述。这些配置为快速训练和高效计算提供了保障，使得模型能够在较短时间内完成训练过程。

所有输入图像均经过裁剪处理，调整至512像素×512像素，以满足UNet网络输入的需求。为了增强模型的泛化能力，数据集还进行了多种数据增强技术的应用，这些技术模拟了不同的拍摄条件和环境变化，从而提升了模型在实际应用中的鲁棒性和适应性。

在模型训练过程中，采用Adam优化器作为主要优化算法，初始学习率设定为0.000 01。选择Adam优化器的原因在于其自适应调整学习率的能力，能够根据梯度的一阶和二阶矩动态更新学习率，从而加快收敛速度并提高模型性能。经过多次实验验证，一阶矩估计的指数衰减率设置为0.9，二阶矩估计的衰减率设置为0.99，这些参数的选择有效地提升了模型在训练过程中的稳定性与收敛效率。

2.3 评价指标

本实验采用语义分割通用的交并比IoU，准确率Precision，召回率Recall，$F1$-Score来对实验结果进行对比[13]，验证模型效果，几种评价指标的计算公示如式（6）~（9）所示。

$$\mathrm{IoU}=\frac{TP}{TP+FP+FN} \tag{6}$$

$$\mathrm{Precison}=\frac{TP}{TP+FP} \tag{7}$$

$$\mathrm{Recall}=\frac{TP}{TP+FN} \tag{8}$$

$$\mathrm{F1-Score}=2\times\frac{\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}\times\mathrm{Recall}} \tag{9}$$

式中，TP为预测结果中模型推断类型与该像素真实类型一致的像素点数量；TN为预测结果中该像素真实类型与模型推断类型不一致的像素点数量；FP为预测结果中模型推断类型与该像素真实类型不一致的像素点数量；FN为所有被分类错误像素点的集合。

2.4 实验结果分析

为了有效验证所提出的半监督语义分割模型的性能，本文选择了当前流行的几种语义分割网络作为对比模型，包括FCN[11]、UNet和DeepLabV3。通过将这些对比模型与提出的方法进行比较，可以全面评估新方法的性能和优越性。

在实验中，本文设置了4种标签比例条件进行对比实验，分别为：标签数据在数据集中占比为1/8、标签数据在数据集中占比为1/4、标签数据在数据集中占比为1/2和标签数据在数据集中占比为1，即全监督训练。这些比例对应于不同的标签数据量，旨在模拟真实场景中不同程度的数据标注情况。通过在这些标签比例下进行多次实验，可以更好地了解模型在不同标签量条件下的表现，评估其在实际应用中的有效性和稳定性。实验结果如表1所示。

表1 本文模型与其他深度学习算法对比

网络模型	标签数据比例	无标签数据比例	Precision/%	Recall/%	F1-Score/%	IoU/%
FCN	1/8	7/8	52.53	52.57	52.55	50.17
UNet			52.11	51.25	51.68	51.63
DeeplabV3			51.15	52.96	52.04	50.96
本文算法			62.38	59.15	60.72	60.65
FCN	1/4	3/4	53.24	53.33	53.28	52.61
UNet			60.47	59.36	59.91	58.83
DeeplabV3			59.21	58.81	59.01	58.30
本文算法			74.32	72.11	73.20	72.16
FCN	1/2	1/2	68.75	67.59	68.17	65.47
UNet			66.33	64.29	65.29	64.30
DeeplabV3			74.76	73.89	74.32	73.18
本文算法			90.26	88.14	89.19	89.32
FCN	1	0	83.69	86.23	84.94	82.15
UNet			87.23	86.07	86.65	85. 22
DeeplabV3			86.25	85.64	85.94	85. 24
本文算法			85.94	86.44	86.19	84. 20

从表1可以看出，当标签数据比例为1/8时，各模型的各项指标精度普遍较低。

随着标签比例的增加至1/4时，本文算法的性能有了显著提升。在这一标签比例下，本文算法的精度已超越了FCN和UNet在标签比例为1/2时的表现。这表明，所提出的半监督算法在适量的标签数据下，能够充分发挥其性能，提升分割精度，甚至在标签较多的情况下也能够表现出优越的效果。当标签比例进一步增加到1/2时，本文算法的性能已经超过了其他对比模型在全监督学习条件下的实验结果。

在对Vaihingen数据集进行模型性能比较时，本文还分析了在标签比例为1/2时，各种模型在所有特征上的IoU表现。根据表2的结果，可以看到本文算法在所有类别中的表现均优于FCN、UNet和DeeplabV3。

表2 对比方法在标签比例为1/2各个类别IoU性能结果

标签比例	类别	方法模型			
		FCN	UNet	DeeplabV3	本文算法
1/2	水体	67.3	67.2	68.1	85.2
	建筑	64.9	65.3	69.23	94.56
	植被	64.21	57.35	79.1	84.05
	树木	61.7	57.4	72.34	88.2
	车体	66.48	65.23	72.19	91.88
	杂物	68.23	67.32	78.12	92.03

综上所述，实验证明所提出的半监督语义分割模型在各种标签比例条件下都具有较好的性能，能够有效地减少标签成本，同时保持或提升模型的分割精度。

图 2 展示了各模型在不同标签比例下的分割结果。从图中看出，随着标签比例的增加，本文算法在边缘修复和异常点减少方面的表现越来越出色。在标签数据非常有限的情况下，本文算法仍能够有效预测图像的主要结构信息，并随着标签量的增加进一步提升分割质量。这表明所提出的半监督学习方法在减少标签数据量的同时，能够保持或提高模型的分割性能，体现了其在实际应用中的潜力和优势。尤其在标注资源有限的情况下，为遥感图像分割提供了一种有效的解决方案。

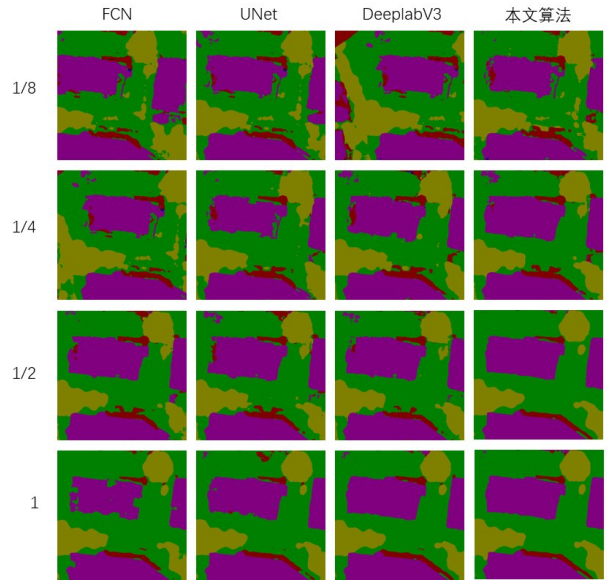


图 2　测试图像在各模型算法在不同标签比例下的分割结果

3　结　语

本研究针对遥感图像分割中标注数据依赖性强、训练成本高的问题，提出了一种基于交叉验证的半监督学习方法。实验结果表明，本方法通过协同利用少量标注数据与大量未标注数据，在 Vaihingen 数据集上仅需 1/2 标注量即可超越全监督基准模型性能（IoU 89.32%，准确率 90.26%），验证了基于一致性假设的交叉监督机制在挖掘未标注数据特征方面的有效性。这一突破显著降低了遥感图像分割的标注成本，为实际应用中的大规模数据场景提供了可行解决方案。

参考文献

[1] 孙航,方帅领,但志平,等. 层级特征交互与增强感受野双分支遥感图像去雾网络[J]. 遥感学报,2023,27(12):2831-2846

[2] Kliestik T, Musa H, Machova V, et al. Remote Sensing Data Fusion Techniques, Autonomous Vehicle Driving Perception Algorithms, and Mobility Simulation Tools in Smart Transportation Systems [J]. Contemporary Readings in Law and Social Justice,2022,14(1):137-152

[3] Chen Q, Li H, Lu G. Training Esrgan with Multi-scale Attention U-Net Discriminator [J]. Scientific Reports, 2024, 14 (1):29036

[4] 肖泽标,梁杰文. 基于改进DeepLabV3+的高分辨率卫星影像分割方法[J]. 测绘标准化,2024,40(3):50-57

[5] Tarvainen A, Valpola H. Mean Teachers are Better Role Models: Weight-averaged Consistency Targets Improve Semi-supervised Learning Results[J].Advances in Neural Information Processing Systems,2017(10):1195-1204

[6] Sohn K, Berthelot D, Carlini N, et al. Fix Match: Simplifying Semi-supervised Learning with Consistency and Confidence[J]. Advances in Neural Information Processing Systems,2020,33(12):596-608

[7] 陈杰,王宁,李想,等. 多模态数据增强的遥感影像半监督分割网络 Semi-remoteNet[J]. 遥感技术与应用,2023,38 (2): 345-356

[8] 王俊迪,赵慧,罗耀华,等. 深度学习在植被高光谱遥感中的研究综述[J]. 遥感信息,2024,39(4):11-22

[9] 成浩维,资文杰,彭双. 基于半监督学习的三维 Mesh 建筑物立面提取与语义分割方法[J]. 郑州大学学报(理学版),2023,55(4):8-15

[10] 冯权泷,陈泊安,李国庆. 遥感影像样本数据集研究综述[J]. 遥感学报,2022,26(4):589-605

[11] Liang P,Chen J,Wu Y, et al. Data Free Knowledge Distillation with Feature Synthesis and Spatial Consistency for Image Analysis [J]. Scientific Reports,2024,14(1):27557

[12] 单飞龙,吕鹏远,李梦晨. 基于CNN-transformer半监督交叉学习的遥感图像场景分类方法[J].宁夏大学学报(自然科学版),2024,45(3):325-332

[13] 王永财,万华伟,高吉喜,等. 基于深度学习语义分割模型的草地植被盖度估算对比研究[J]. 环境科学研究,2024,37 (10): 2299-2309

[14] 向儒萱毅. 基于深度神经网络的海陆中尺度对流系统多光谱遥感图像识别[D]. 南京.南京信息工程大学,2024