

# 基于弱监督与光谱指数的哨兵2号图像裸土提取

赵焱海

天津工业大学物理科学与技术学院, 天津

收稿日期: 2023年12月7日; 录用日期: 2024年1月1日; 发布日期: 2024年1月8日

## 摘要

利用哨兵2号图像进行裸土区动态制图对环境管理和生态恢复具有重要意义。深度学习技术已经彻底改变了土地利用与土地覆盖分类的方法, 包括裸露土地区域的映射。然而, 目前的深度学习方法存在两个主要问题, 即标记成本高和模型性能较差的问题。在本文中, 我们开发了一种新的深度语义分割网络 O-Net 来解决当前的问题。O-Net 具有典型的编码器-解码器结构, 其中编码器和解码器都可以实例化为一个特定的全卷积网络, 其中编码器用于预测裸土面积的提取, 解码器用于重建输入图像斑块的裸土指数。我们基于弱监督(如不完全或不准确的标签)来训练网络参数。在我们的注释浙江数据集上的实验表明, 所提出的方法可以实现较先进的性能。

## 关键词

裸土提取, 多光谱影像, 深度学习, 弱监督, 裸土指数

# Bare Soil Extraction from Sentinel-2 Images with Weakly Supervised and Spectral Index

Yihai Zhao

School of Physical Science and Technology, Tianjin Polytechnic University, Tianjin

Received: Dec. 7<sup>th</sup>, 2023; accepted: Jan. 1<sup>st</sup>, 2024; published: Jan. 8<sup>th</sup>, 2024

## Abstract

Dynamic mapping of bare soil areas using sentinel-2 images is important for environmental management and ecological restoration. Deep learning technology has revolutionized the approaches of Land Use and Land Cover classification including the mapping of bare soil areas. However, there are two main problems with current deep learning methods, namely the high cost of labeling and poor model performance. In this paper, we develop a new deep semantic segmentation network called O-Net for tackling the current problems. The O-Net has a typical encoder-decoder structure,

among which both the encoder and the decoder can be instantiated as a specific fully convolutional network, where the encoder is used to predict the extraction of bare soil areas and the decoder is used to reconstruct the bare soil indices of the input image patches. We train the network parameters based on weak supervision such as incomplete or inaccurate annotations. Experiments over our annotated Zhejiang Dataset demonstrate that the proposed method can achieve more advanced performance.

## Keywords

Bare Soil Extraction, Multispectral Images, Deep Learning, Weak Supervision, Bare Soil Index

Copyright © 2024 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. 引言

随着城市化进程的加快，人类和自然活动，如工程建设、森林采伐、土地沙漠化、泥沙淤积等，往往造成大量裸土，造成严重的生态安全风险，如粉尘污染、土壤侵蚀等[1]。

随着时代的发展，遥感技术由于可以覆盖广阔的地理区域，对大面积的地表进行连续监测，因此逐渐被广泛的投入使用。同时通过遥感影像，可以迅速获取大范围内的环境信息，有效监控森林砍伐、土地荒漠化的进程，同时裸土遥感影像的自动提取可以为环境管理和生态恢复提供了基础地理数据。像哨兵 2 号这样的多波段图像由于具有重访周期短的优点，适用于包括裸土在内的土地利用和土地覆盖 (LULC) 的高频监测。

然而遥感图像通常具有大规模、高维度的特征，例如卫星图像、航拍图像等，这些图像往往需要进行像素级的标注才能用于监督学习。然而，大规模的像素级标注成本高昂，而且需要专业知识和人力资源。而现在流行的弱监督方法可以利用较为简单、低成本的标注信息，如不确切标注、不准确的标注或不完全标注的数据，来进行学习和训练，从而一定程度上克服了数据标注困难的问题。

因此为了降低裸土提取的标注成本，充分利用光谱信息等弱监督信息，一定程度提高模型的裸土提取性能，我们提出了一种新的深度学习网络，称为 O-Net，用于从 sentinel-2 图像中提取裸露土地区域。该网络被设计成一个编码器 - 解码器结构，其中编码器输出预测结果，解码器输出重建的 BI 裸土指数，这些重建的 BI 裸土指数寻求与提取出的裸土光谱指数的特征对齐。本文的主要贡献包括：

(1) 我们创建了一个 sentinel-2 图像数据集，由一个不完全标注的训练集和一个完全标注的测试集组成，该数据集可用于训练和测试弱监督模型。

(2) 我们提出了一种新的编码器 - 解码器网络结构，称为 O-Net，其中编码器和解码器都使用特定的全卷积网络。编码器输出语义分割结果，解码器输出 BI 裸土指数重构结果。

我们在创建的数据集上进行了实验，结果表明所提出的 O-Net 网络可以从不完全或不准确的监督中获得学习，并且所获得的 O-Net 模型在裸土提取的测试中表现良好。

## 2. 相关工作

通过遥感影像进行裸土提取的典型方法包括三类：光谱指数方法、机器学习方法和深度学习方法。

光谱指数法方面，基于多波段的反射率差异，不同的学者分别设计了裸土指数(BI) [2]、改良裸土指

数(MBI) [3]、干裸土指数(DBSI) [4]等光谱指数方法, 利用一定的统计阈值提取裸土区域。

传统的机器学习方法依靠像素级标注来训练分类器进行裸土提取, 典型的包括 k 近邻分类器[5]、人工神经网络[6]、决策树[7]、支持向量机[8]、XGBoost [9]等, 但上述方法主要依据浅层光谱特征进行分类, 无法有效提取时空上下文信息等深层次特征, 且特征的选择靠经验为主。

随着人工智能技术的发展, 深度学习因其深层特征表征能力、端到端的训练优势而逐渐成为主要的研究趋势, 同时也为计算机视觉和遥感图像分析提供了更好的选择。其中, 针对图像场景分类的代表性的研究包括: Lecun 等人[10]提出的 LeNet 作为一种卷积神经网络, 被广泛的用于数字识别; Krizhevsky 等人[11]提出的 AlexNet 验证了使用纯监督学习, 一个大型的深度卷积神经网络能够在高度具有挑战性的数据集上实现特别好的结果; Simonyan 等人[12]提出的 VGG 评估了非常深的卷积网络可用于大规模的图像分类, 表明深度有利于分类精度, 实现先进的性能; He 等人[13]提出的 ResNet 证明了残差网络更容易优化, 并可以从相当大的深度网络获得较高的准确性, 且不会出现退化现象。针对图像语义分割的代表性研究包括: Long 等人[14]提出的全卷积网络 FCN, 它将来自较深的粗层的语义信息与来自较浅的细层的外观信息结合起来, 以生成精确而详细的分割结果; Ronneberger 等人[15]提出的 U-Net 网络, 证明了通过依赖于数据增强的强大使用, 可以更有效地使用可用的标注样本, 从而可以从很少的图像中进行端到端训练得到很好的结果; Chen 等人[16]提出了 DeepLabv3+来扩大接受野。这些全卷积骨干网络已在 LULC 分类中得到了应用。例如, Liu 等人[17]将 CBAM 模块(卷积块注意模块)和 SE 模块(挤压激励模块)结合到 DeepLabv3+网络中, 增强了不透水表面的提取能力。Yao 等人[18]提出了一种基于 swin-unet [19]的分割网络, 用于 Sentinel-2 多光谱图像的裸土提取。

深度学习方法的一个主要瓶颈是需要大量的标注数据, 这在中分辨率遥感领域尤为昂贵。有鉴于此, 弱监督的方法得到了越来越多的关注。弱监督方法通常采用三种主要策略: 不准确监督、不准确监督和不完全监督[20]。例如, Schmitt 等人[21]提出了一种基于不准确监督的方法, 该方法通过自动上采样低分辨率的 MODIS LCCS (土地覆盖分类系统) 土地利用图来注释哨兵图像。实验表明, 与原始低分辨率地图相比, 导出的模型可以检索到更多的细节。Nivaggioli 等人[22]利用图像级不精确监督和像素间的邻域关系自动生成像素级标注, 推导的深度语义分割模型与实际的像素级标注性能相同。Qiao 等人[23]也提出了一种基于图像级注释的方法, 用于从震后高分辨率卫星图像中提取受损建筑, 该方法通过提高类激活图的质量来增强模型性能。Lian 等人[24]提出了一种基于不完全点标注的弱监督道路分割方法, 该方法有效地将深度卷积神经网络的特征表示能力与启发式算法的可解释性相结合。Zhang 等人[25]利用图像级标签来生成可靠的像素级注释, 并设计了一个完全端到端的网络来学习预测分割图, 同时考虑浅层和高层特征, 在 VOC 数据集上取得了更好的结果。Rong 等人[26]提出使用两个交互式网络的协同训练范例来改进不确定像素的学习, 并提出了一种边界增强策略来增强对困难边界区域的预测, 从而实现了较好的分割性能。Zhou 等人[27]通过在用于网络学习和推理的大量弱标记训练数据中协同探索丰富的语义上下文, 从而实现了更全面的语义聚合, 获得了强大的细粒度语义分割能力。Du 等人[28]提出了弱监督的像素与原型对比的方法, 可以提供像素级的监督信号来缩小差距, 并在 PASCAL VOC 2012 数据集上取得了先进水平。Zhu 等人[29]提出了一种有效的两阶段弱监督语义分割方法, 通过对特征进行聚类以获得子类别伪标签, 然后设计了一个自注意力模块来进一步提高响应图的质量。在多个数据集上获得了较好的性能和优越性。

### 3. 数据集

#### 3.1. 数据集的来源

本研究使用的数据集是由我们手工标注的浙江数据集。浙江数据集的地理位置如图 1 所示。其中,

训练集提取自中国浙江省杭州市(蓝色区域), 测试集提取自中国浙江省绍兴市(黄色区域)。浙江数据集的训练集采用人工不完全标注方式标注用于弱监督训练的研究, 测试集采用完全标注方式标注用来更好的测试裸土提取的效果。

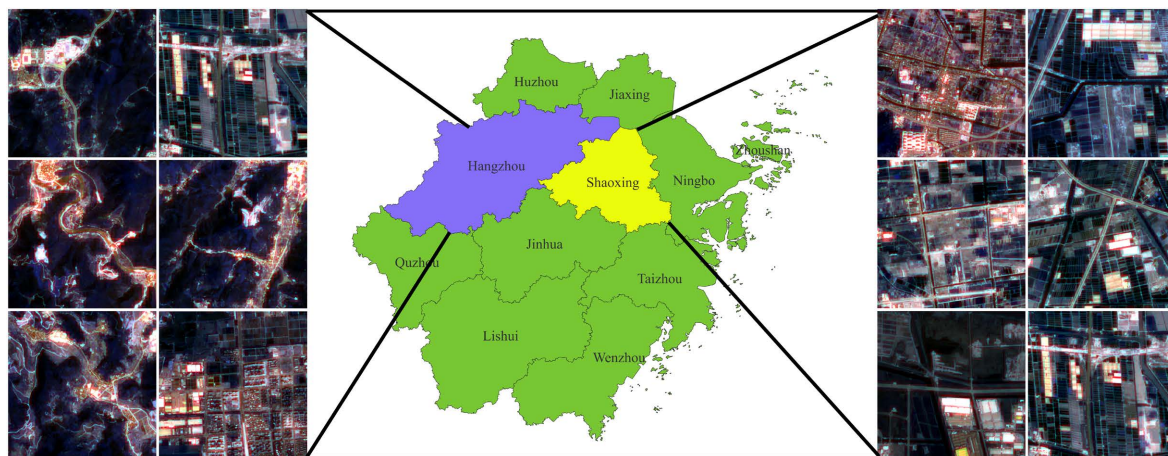


Figure 1. Data sources

图 1. 数据来源

### 3.2. 数据集的构建

浙江省数据集的创建过程如下。首先, 从欧洲航天局(ESA)网站下载研究区域的 Sentinel-2 光学遥感图像, 如图 1 所示。其次, 利用 SNAP 软件中的超分辨率插件, 对 Sentinel-2 多光谱图像均匀上采样至 10 米的空间分辨率。最后, 选取了 832 个包含裸土区域的图像斑块, 每个斑块的大小为  $256 \times 256$  像素。这些斑块被分成两组: 来自杭州市的 780 块作为训练集, 来自绍兴市的 52 块作为测试集。训练集的标注像素总数为 332,447 个, 测试集的标注像素总数为 694,798 个, 如表 1 所示。

Table 1. The details of the Zhejiang Dataset

表 1. 浙江数据集详情

	训练数据	测试数据	总的的数据
裸土斑块数量(块)	780	52	832
裸土像素数量(个)	332447	694798	102745

## 4. 本文的方法

### 4.1. O-Net 网络架构

我们开发了一种称为 O-Net 的语义分割网络结构, 专门为弱监督训练量身定制。如图 2 所示, O-Net 被设计为编码器 - 解码器网络结构, 其中编码器和解码器都可以利用特定的全卷积网络。在本文中, 我们将标准 Deeplabv3+网络作为编码器, 将标准 U-Net 网络作为解码器。编码器以原始图像为输入, 通过前向传播输出语义分割预测, 解码器以语义分割预测为输入, 通过前向传播输出重构的 BI 指数。

这个设计基于两个考虑。一方面, 解码器可以被视为在语义分割期间特定的正则化项, 以减轻可能由不完全的注释和小样本训练引起的过拟合。另一方面, 我们假设 BI 指数可能包含与裸露土地区域的一定相关性, 这可能有助于提高语义分割模型提取裸土的能力。

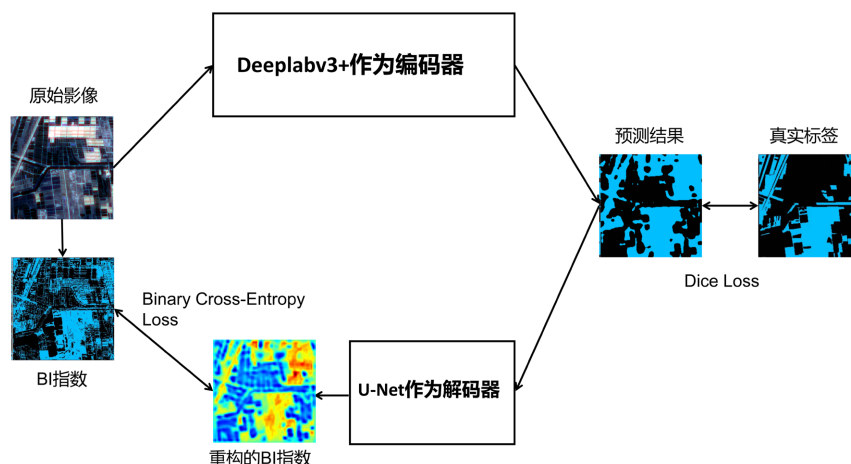


Figure 2. The O-Net network structure  
图 2. O-Net 网络架构

## 4.2. 编码器模块

在本研究中，我们选用了 Deeplabv3+作为编码器的主干网络。Deeplabv3+是 DeepLab 系列的最新语义分割模型。该网络通过使用 ASPP 模块和深度可分离卷积(DSC)来融合多尺度特征，平衡了精度和时间消耗。在编码器中，骨干网络对图像处理的结果分为两部分：一个作为浅层特征直接引入解码器，另一个在不同尺度的萎缩卷积下，通过并行 ASPP 模块获得特征提取。通过使用  $1 \times 1$  卷积层对特征进行压缩，将不同尺度的结果合并为深度特征。然后，将深层特征上采样到解码器中。在解码器中，浅层特征和深层特征可以被连接成合并的特征图，随后，卷积层和上采样层对合并的特征表进行处理，最终获得预测结果。这种编码器提取的特征可以保留原始图像的语义信息，从而提高了遥感影像的裸土区域识别精度。

## 4.3. 解码器模块

本研究解码器模块采用了 U-Net 的解码器部分，因为 U-Net 不仅可以捕捉全局信息，还可以利用编码器部分提供的特征进行逐层重建，从而更好地还原图像细节。这种机制使得 U-Net 非常适合用于解码器，可以有效地恢复图像的细节，用来重构裸土指数。裸土指数也属于光谱指数的一种，光谱指数是遥感图像处理中常用的一种指标，通过计算不同波段之间的比值或差值，能够反映出地表覆盖物在不同波段上的特征差异，从而实现地物分类、植被指数计算等应用。光谱指数的意义在于可以提取出遥感图像中的有效信息，对于遥感图像的解译和分析具有重要作用。因此，光谱指数在遥感图像处理中拥有着广泛的应用前景及重要意义。由于我们进行的是裸土的二分类识别，所以进行单个类别的裸土指数重构效果更佳。这里的指数采用了裸土指数 BI。我们认为通过计算并且重构裸土指数，可以抓住 BI 指数与裸露土地区域的一定相关性，从而提高裸土提取的性能。

## 4.4. 损失函数

我们将 O-Net 网络编码器部分输出的特征与局部标签数据计算其损失值为 Dice Loss；将 O-Net 网络解码器部分输出的裸土指数特征与计算的 BI 指数计算其损失为 Binary Cross-Entropy Loss；基于编码-解码的重构网络总损失值为 Loss，以总损失值最小化为目标训练网络模型。具体的，我们总的损失函数表达为：

$$\text{Loss} = (\text{Dice Loss}) + (\text{Binary Cross-Entropy Loss})$$

## 4.5. 训练方法

我们选择 RMSprop 优化器作为优化器，旨在通过最小化损失函数来调整模型的权重和偏差。具体来说，我们将初始学习率设为  $1e-3$ ，权重衰减系数设为  $1e-8$ ，动量因子设为 0.9。这样的参数配置有效地更新和控制梯度，加速模型的收敛。我们在训练阶段将批大小设置为 16。我们使用 PyTorch 编程框架实现，并使用一台机器的 NVIDIA GeForce RTX 3080 GPU 来加速训练过程。

## 5. 实验结果

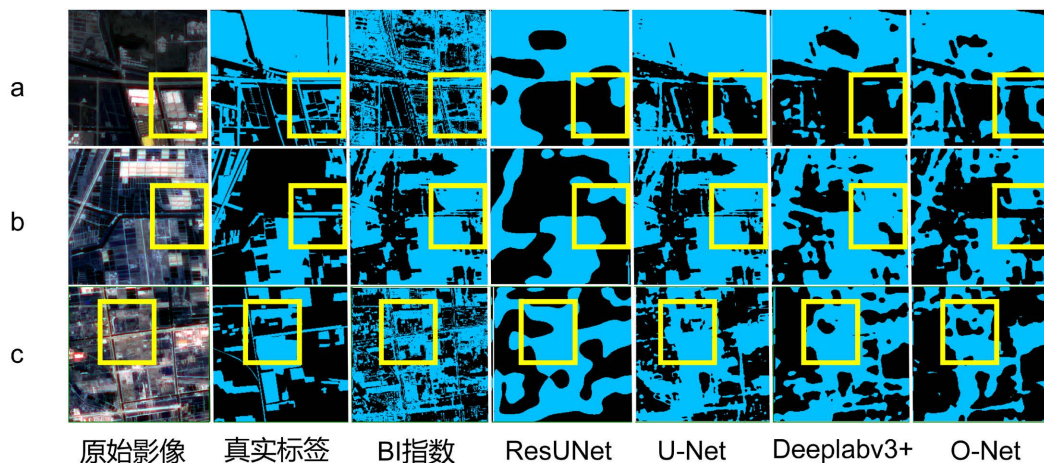
为了确保实验结果的可比性，我们将我们提出的分割模型与现有的基线方法进行了对比，使用了相同的训练和测试样本。我们采用了 IoU 和 F1 指数来评估模型的性能，并将实验结果展示在表 2 中。通过对比我们可以发现，我们的实验结果整体上超过了大多数现有的基线方法，比最好的 Deeplabv3+方法在 F1 指数方面提升了 5.28%，在 IoU 方面提升了 3.27%。

**Table 2.** Test results over the Zhejiang Dataset

**表 2.** 浙江数据集的测试结果

实验方法	F1	IoU
BI 指数法	0.4473	0.2881
ResUNet	0.5487	0.3768
U-Net	0.5673	0.3897
Deeplabv3+	0.5743	0.4034
O-Net	<b>0.6271</b>	<b>0.4361</b>

为了更好的展现不同方法提取裸土的结果，我们把不同方法的提取结果进行了可视化，如图 3 所示，其中 a, b, c 分别代表不同的图像区域，图中蓝色代表裸土区域，黑色代表非裸土区域，黄色为重点关注的裸土区域，可以看出 O-Net 网络能够提取出较多的有效裸土区域，具有较好的裸土提取性能。



**Figure 3.** Comparison of the original image with the extraction results of five different methods

**图 3.** 原始图像与五种不同方法的提取效果图对比

为了更好地解释 O-Net 网络为什么相较于本实验中的其他方法相比具有更好的性能，我们使用了特征的二维空间投影(t-SNE)分别对 O-Net 网络与较好的 Deeplabv3+网络进行二维空间投影可视化呈现进行对比分析。特征二维空间投影的意义在于将特征从高维特征空间映射到二维平面，从而使特征之间的关

系更加直观和容易理解。也可以帮助我们更直观地理解特征之间的关系，如是否存在线性相关、聚集等。具体的二维空间投影可视化结果如图 4 所示，其中紫色代表裸土类，黄色代表非裸土类，从图中我们可以看出 O-Net 网络对于裸土类和非裸土类类别的可分性更强，意味着它们在特征空间中的表示区分度更高，有着更好的语义信息区分度，从而可以有效的提取出更多的有效裸土区域。

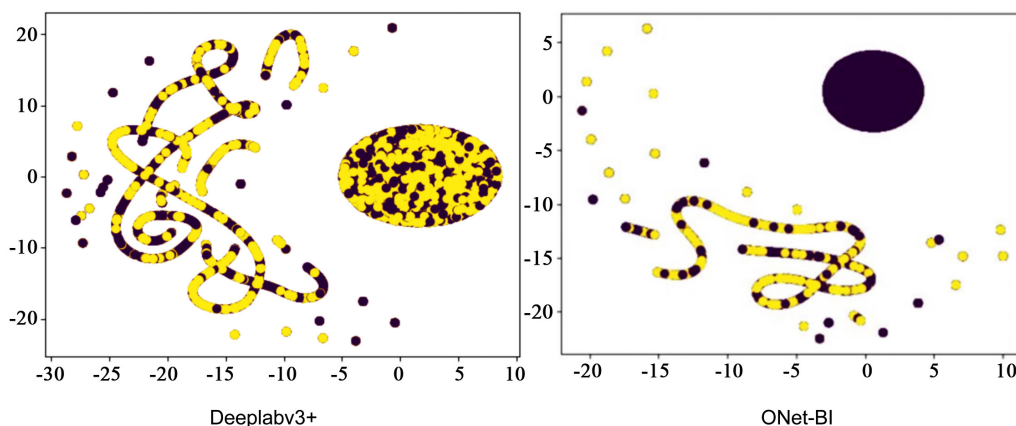


Figure 4. 2D spatial projection of output features  
图 4. 输出特征的二维空间投影

## 6. 结论

总的来说，通过适当利用裸土指数这一不准确的弱监督信息，我们提出了一种新的深度语义分割网络 O-Net。实验结果表明，该方法在浙江数据集上是有效的。我们可以发现我们的实验结果与基本的理论假设较为吻合。具体而言，基于上述实验结果，我们可以发现 O-Net 可以有效地利用裸土指数增强的弱监督信息，并且这种信息的利用对提高在不完全弱监督的情况下的裸土提取能力有很大意义。在未来，我们将进一步研究并引入一些新的具有较强地物信息的光谱指数。例如，整合 MNDWI(归一化差异水指数)可能有助于缓解土壤湿度的负面影响，或者整合 NDBI (归一化差异建筑指数)可能有助于缓解建筑物对于裸图提取的混淆影响。同时在未来，我们还将进一步研究基于所提出的 O-Net 结构在跨不同区域场景中的裸土提取效果。

## 参考文献

- [1] Xu, H. (2013) Dynamics of Bare Soil in a Typical Reddish Soil Loss Region of Southern China: Changting County, Fujian Province. *Scientia Geographica Sinica*, **33**, 489-496.
- [2] Chen, W., Liu, L., Zhang, C., et al. (2004) Monitoring the Seasonal Bare Soil Areas in Beijing Using Multitemporal TM Images. *2004 IEEE International Geoscience and Remote Sensing Symposium*, Vol. 5, 3379-3382.
- [3] Nguyen, C.T., Chidthaisong, A., Kieu Diem, P., et al. (2021) A Modified Bare Soil Index to Identify Bare Land Features during Agricultural Fallow-Period in Southeast Asia Using Landsat 8. *Land*, **10**, Article No. 231. <https://doi.org/10.3390/land10030231>
- [4] Rasul, A., Balzter, H., Ibrahim, G.R.F., et al. (2018) Applying Built-Up and Bare-Soil Indices from Landsat 8 to Cities in Dry Climates. *Land*, **7**, Article No. 81. <https://doi.org/10.3390/land7030081>
- [5] Zhu, H. and Basir, O. (2005) An Adaptive Fuzzy Evidential Nearest Neighbor Formulation for Classifying Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, 1874-1889. <https://doi.org/10.1109/TGRS.2005.848706>
- [6] Kavzoglu, T. and Mather, P.M. (2003) The Use of Backpropagating Artificial Neural Networks in Land Cover Classification. *International Journal of Remote Sensing*, **24**, 4907-4938. <https://doi.org/10.1080/0143116031000114851>
- [7] Friedl, M.A. and Brodley, C.E. (1997) Decision Tree Classification of Land Cover from Remotely Sensed Data. *Re-*

- mote Sensing of Environment*, **61**, 399-409. [https://doi.org/10.1016/S0034-4257\(97\)00049-7](https://doi.org/10.1016/S0034-4257(97)00049-7)
- [8] Gualtieri, J.A. and Cromp, R.F. (1999) Support Vector Machines for Hyperspectral Remote Sensing Classification. *27th AIPR Workshop: Advances in Computer-Assisted Recognition*, Vol. 3584, 221-232. <https://doi.org/10.1117/12.339824>
- [9] Chen, T., He, T., Benesty, M., *et al.* (2015) Xgboost: Extreme Gradient Boosting. R Package Version 0.4-2, 1-4.
- [10] LeCun, Y., Bottou, L., Bengio, Y., *et al.* (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. <https://doi.org/10.1109/5.726791>
- [11] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90.
- [12] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition.
- [13] He, K., Zhang, X., Ren, S., *et al.* (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [14] 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>
- [15] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference*, Munich, 5-9 October 2015, 234-241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [16] 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](https://doi.org/10.1007/978-3-030-01234-2_49)
- [17] Liu, J., Zhang, Y., Liu, C., *et al.* (2023) Monitoring Impervious Surface Area Dynamics in Urban Areas Using Sentinel-2 Data and Improved Deeplabv3+ Model: A Case Study of Jinan City, China. *Remote Sensing*, **15**, Article No. 1976. <https://doi.org/10.3390/rs15081976>
- [18] Yao, J. and Jin, S. (2022) Multi-Category Segmentation of Sentinel-2 Images Based on the Swin UNet Method. *Remote Sensing*, **14**, Article No. 3382. <https://doi.org/10.3390/rs14143382>
- [19] Cao, H., Wang, Y., Chen, J., *et al.* (2022) Swin-unet: Unet-Like Pure Transformer for Medical Image Segmentation. *European Conference on Computer Vision*, Tel Aviv, 23-27 October 2022, 205-218. [https://doi.org/10.1007/978-3-031-25066-8\\_9](https://doi.org/10.1007/978-3-031-25066-8_9)
- [20] Zhou, Z.H. (2018) A Brief Introduction to Weakly Supervised Learning. *National Science Review*, **5**, 44-53. <https://doi.org/10.1093/nsr/nwx106>
- [21] Schmitt, M., Hughes, L.H., Qiu, C., *et al.* (2019) SEN12MS—A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion. <https://doi.org/10.5194/isprs-annals-IV-2-W7-153-2019>
- [22] Nivaggioli, A. and Randrianarivo, H. (2019) Weakly Supervised Semantic Segmentation of Satellite Images. 2019 *Joint Urban Remote Sensing Event (JURSE) IEEE*, Vannes, 22-24 May 2019, 1-4. <https://doi.org/10.1109/JURSE.2019.8809060>
- [23] Qiao, W., Shen, L., Wang, J., *et al.* (2023) A Weakly Supervised Semantic Segmentation Approach for Damaged Building Extraction from Post-Earthquake High-Resolution Remote-Sensing Images. *IEEE Geoscience and Remote Sensing Letters*, **20**, 1-5. <https://doi.org/10.1109/LGRS.2023.3243575>
- [24] Wang, J., Shao, Z., Huang, X., *et al.* (2022) From Artifact Removal to Super-Resolution. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, 1-15. <https://doi.org/10.1109/TGRS.2022.3196709>
- [25] Zhang, B., Xiao, J., Wei, Y., *et al.* (2022) End-to-End Weakly Supervised Semantic Segmentation with Reliable Region Mining. *Pattern Recognition*, **128**, Article ID: 108663. <https://doi.org/10.1016/j.patcog.2022.108663>
- [26] Rong, S., Tu, B., Wang, Z., *et al.* (2023) Boundary-Enhanced Co-Training for Weakly Supervised Semantic Segmentation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, 17-24 June 2023, 19574-19584. <https://doi.org/10.1109/CVPR52729.2023.01875>
- [27] Zhou, T., Zhang, M., Zhao, F., *et al.* (2022) Regional Semantic Contrast and Aggregation for Weakly Supervised Semantic Segmentation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, 18-24 June 2022, 4299-4309. <https://doi.org/10.1109/CVPR52688.2022.00426>
- [28] Du, Y., Fu, Z., Liu, Q., *et al.* (2022) Weakly Supervised Semantic Segmentation by Pixel-to-Prototype Contrast. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, 18-24 June 2022, 4320-4329. <https://doi.org/10.1109/CVPR52688.2022.00428>



- 
- [29] Zhu, H., Geng, T., Wang, J., *et al.* (2023) Improved Sub-Category Exploration and Attention Hybrid Network for Weakly Supervised Semantic Segmentation. *Neural Computing and Applications*, **35**, 10573-10587.  
<https://doi.org/10.1007/s00521-023-08250-4>