

景维鹏,王健,张文钧,等.基于缩影的多时相遥感语义变化检测方法[J].华中农业大学学报,2023,42(3):123-132.
DOI:10.13300/j.cnki.hnlkxb.2023.03.015

基于缩影的多时相遥感语义变化检测方法

景维鹏¹,王健¹,张文钧¹,谷俊涛²,陈广胜¹

1.东北林业大学信息与计算机工程学院,哈尔滨 150040; 2.黑龙江省网络空间研究中心,哈尔滨 150090

摘要 针对高分辨率遥感图像标签稀缺和标签技术增长缓慢限制了多时相语义变化检测发展的问题,提出了采用有噪声、低分辨率的弱标签替代高分辨率标签进行多时相语义变化检测的方法。首先,采用低分辨率卫星数据平滑高分辨率遥感图像输入的质量差异。其次,通过将缩影(epitomes)模型和标签超分辨率算法作为统计推理算法相结合的方法预估高分辨率遥感图像分类图,并拟合一个小型FCN网络对生成的遥感图像分类图进行后处理来改善其分类的效果。最后,通过对比不同时相土地覆盖分类图像之间的差异得出变化检测结果。结果表明,本研究提出的方法与其他多时相语义变化检测方法FCN/all相比,平均交并比(mIoU)提高了8.9个百分点,能够有效检测土地覆盖分类变化。

关键词 弱监督; 标签超分辨率; 缩影; 土地覆盖变化图; 语义分割; 遥感影像; 变化检测

中图分类号 TP79 **文献标识码** A **文章编号** 1000-2421(2023)03-0123-10

随着对地观测技术的不断进步^[1],遥感图像在地面分辨率、空间分辨率等方面都有了巨大的提高,这种提高使得利用遥感图像能够更好地获取地物信息。遥感图像变化检测^[2]是指利用多时相遥感影像(即不同时期覆盖同一地表区域的遥感图像)进行数据处理和影像对比,从而获取该区域地物变化信息的过程。国内外学者针对遥感图像变化检测技术提出了许多方法,如图像差分法^[3]、变化向量分析法^[4]、面向对象法^[5]和基于深度学习的方法^[6-7]等。但是这些方法关注的主要问题是某一区域是否发生变化,属于二分类问题,缺少涉及地物变化的具体类别信息^[8]。而语义分割恰恰可以增加涉及地物变化的具体类别信息^[9]。语义分割可以将不同类别标签赋予高分辨率遥感图像的不同像素,使其可以分析及处理遥感图像中的不同细节。多时相语义变换检测任务^[10]将变换检测和语义分割结合起来,在变与不变的基础上引入语义信息。该任务由于关注了遥感图像从一种地物类型变化为另一种地物类型的信息,可以更好地分析土地利用情况和土地覆盖变化,检测农业用地的扩张程度,管理农业资源^[11-12]。

影响多时相语义变化检测任务发展的主要因素是高空间分辨率和高时间分辨率传感器的激增无法完全与同样丰富的标签相匹配^[13-14],导致高分辨率遥感土地分类标签稀缺。在遥感领域,低分辨率卫星图像和低分辨率土地覆盖标签是定时更新的,并且其图像和标签会公开提供。因此,为了降低高分辨率标签数据的制作成本,从稀疏信息的弱标签中学习来执行标签超分辨率以及进行准确的遥感语义变化检测是一个可行的方案^[15]。解决弱监督分割问题不仅可以依赖于图像级别的注释^[16-17],也可以依赖于稀疏的标签注释。目前已经研究出多种依赖于稀疏标签的半监督语义分割方法^[18]。Robinson等^[19]提出了一种基于深度学习的高分辨率航空影像土地制图的多分辨率数据融合方法。Malkin等^[20]提出了一种“标签超分辨率”技术,该技术使用在不同低分辨率类别标记的粗像素块内出现高分辨率标签的统计数据,通过粗类对网络进行端到端的训练,可以得到一个能够直接预测各个像素的高分辨率标签的模型。Zhao等^[21]通过研究跨地理区域的类外观之间的差异激发了先前关于弱监督分割(标签超分辨率)的替代损失函数的工作。

收稿日期: 2022-08-02

基金项目:国家自然科学基金项目(32171777);黑龙江省应用技术与开发计划项目(GA20A301)

景维鹏, E-mail: jwp@nefu.edu.cn

通信作者: 陈广胜, E-mail: kjc_chen@163.com

但是上述深度卷积网络方法也存在一些缺陷。如深度卷积网络模型很难解释,这种缺陷使得该模型很难通过向体系中添加常识性先验知识来进行改进。而且促进深度卷积网络在图像中的远距离推理架构上具有一定的局限性,由于深度卷积网络的有效感受野通常只有十几个像素,相对来说较小^[22],这会超出远距离的像素对深层神经元起作用的距离。基于以上问题,Jojić等^[23]提出的基于缩影表示的模型通过执行最邻近方法有效地将小图像块用于视觉任务,该方法可以改善深度卷积网络解释性差以及有效感受野小的问题。Malkin等^[24]推导出标签超分辨率算法作为缩影表示的统计推理算法,可以在弱监督下进行语义分割和标签超分辨率。试验结果表明,基于缩影表示的模型因为具有对图像块进行建模的特点,解决了模型解释困难和模型感受野区域狭小的问题,所以该方法可以将先验知识放入到模型中以替代深度卷积网络方法。

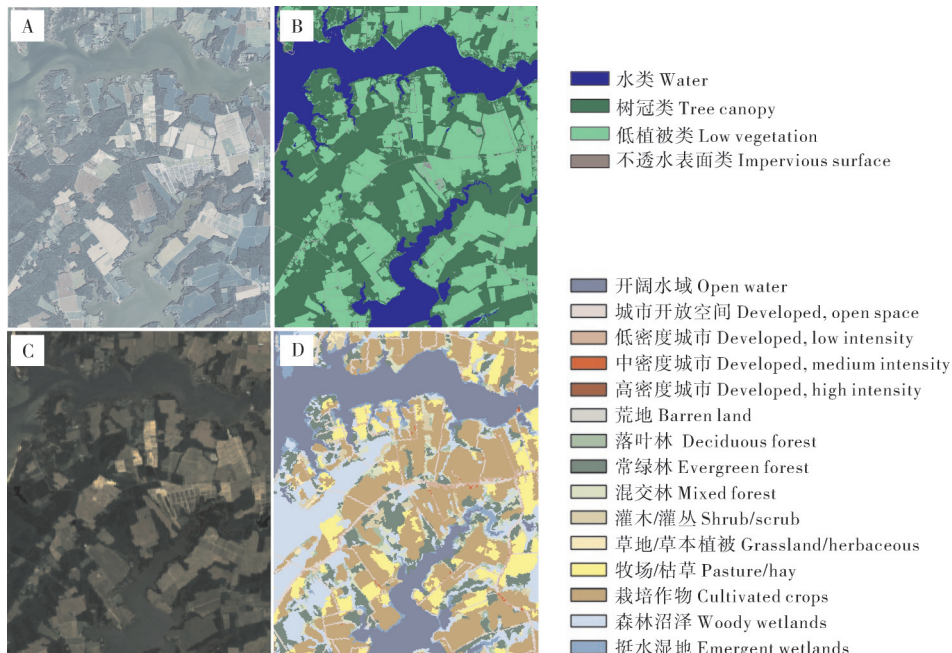
综上所述,鉴于缩影模型在大型的数据集上表现出的性能比较好,并且基于缩影表示的标签超分辨率方法来进行多时相遥感语义变化检测具备可行性。本研究提出了一种基于缩影的标签超分辨率与遥感图像分类标签后处理相结合的多时相遥感语义

变换检测方法,通过弱监督策略,充分利用多源和多时相数据生成具有不匹配噪声标签的高精度和高分辨率土地覆盖变化图,旨在为农业土地覆盖变化检测等提供参考。

1 材料与方法

1.1 试验数据

本研究使用的数据为2013年至2014年美国东北部切萨皮克湾流域的相关数据,图1为数据集示例,其中包括:(1)美国农业部国家农业图像计划(national agriculture imagery program, NAIP)的1 m 高分辨率航空影像,其中包括R、G、B和近红外波段(near infrared, NIR)。(2)美国地质调查局的Landsat 8卫星拍摄的30 m 低分辨率多光谱卫星图像,其中Landsat 8图像是由无云期间卫星地表反射图像的中值所创建,包含9个波段。(3)切萨皮克保护区的高分辨率(1 m)土地覆盖标签,包括4个土地覆盖类别:水(water)、树冠(tree canopy)、低矮植被(low vegetation)、不透水表面类(impervious surface)。(4)美国土地覆盖数据库(national land cover data base, NLCD)的30 m 低分辨率16类土地覆盖标签^[25]。



A: NAIP 影像 NAIP imagery; B: 地面真实土地覆盖标签 Ground truth land cover labels; C: Landsat 8 图像 Landsat 8 image; D: NLCD 土地覆盖标签 NLCD land cover label.

图1 数据集示例

Fig.1 Example dataset

1.2 数据增强

1) 低分辨率数据增强。高分辨率遥感图像在生成过程中难免受到噪声的影响导致图像质量下降,而采用大时间范围内平均的低分辨率Landsat 8图像增强高分辨率图像可以提高模型的性能。选取同一区域的Landsat 8图像和NAIP图像,将低分辨率Landsat 8图像重新采样到与高分辨率NAIP图像相同的维度,然后将2个图像源连接起来,得到新的输入图像。如图1所示,Landsat 8数据图层都被重新采样为NAIP数据中使用的1 m分辨率网格。

2) 输入颜色增强。在小的地理区域内,单个像素颜色是土地覆盖分类的一个非常具有预测性的特征。向输入的NAIP图像添加随机颜色增强,对NAIP图像的每个通道的亮度和对比度随机调整5%。对于NAIP图像的单个通道,求出该通道的平均像素强度,采样参数在0.95~1.05随机选择,并对该通道里面的所有像素进行调整。最后,对NAIP图像的所有通道应用此方法,计算出颜色随机变化后的图像。

1.3 标签超分辨率方法

采用基于缩影表示的标签超分辨率方法,缩影是图像的浓缩版本,它仍然包含重建图像所需要的纹理信息和形状基元,可以实现从稀疏标签中学习并执行标签超分辨率算法从而生成高分辨率土地覆盖制图。该方法首先是实现基于缩影模型的弱监督语义分割,再通过标签超分辨率算法生成高分辨率土地覆盖制图。

1) 基于缩影的语义分割。缩影是图像块高斯混合模型的升级版,由图像块 x^t 展开为向量 $x^t = \{x_{i,j,k}^t\}$ 组成的训练集,其中 i, j 是图像块中的坐标, k 是光谱通道(R, G, B, \dots),使用One-hot标签嵌入对应的向量 $y^t = \{y_{i,j,l}^t\}, l \in \{1, \dots, L\}$ 进行计算。在本研究提出的缩影混合模型中,遥感图像、遥感图像标签数据上的分布借助潜在变量 $s \in \{1, \dots, S\}$ 表示,其公式为:

$$p(x^t, y^t) = \sum_{s=1}^S p(x^t|s) p(y^t|s) p(s) \quad (1)$$

其中, $p(s)$ 是混合分量 s 的概率,而条件概率 $p(x^t|s)$ 描述了潜在变量 s 生成的遥感图像块中允许的变化, $p(y^t|s)$ 描述了遥感图像块上可能生成的标签。在本研究的模型下,对1张新遥感图像 x 的期待分割结果 y 的估计为:

$$p(y|x) = \sum_s p(s|x) p(y|s) \quad (2)$$

$p(x|s)$ 的自然选择为对角高斯分布:

$$p(x|s) = \prod_{i,j,k} \frac{e^{-\frac{1}{2}(x_{i,j,k} - \mu_{s,i,j,k})^2 / \sigma_{s,i,j,k}^2}}{(2\pi\sigma_{s,i,j,k}^2)^{\frac{1}{2}}} \quad (3)$$

$p(y|x)$ 是遥感图像标签上分类分布的简单乘积。混合分量 s 的均值包含像素值 $\mu_{s,i,j,k}$,而协方差矩阵用其对角元素 $\sigma_{s,i,j,k}^2$ 表示,即各个像素 i, j 的不同颜色通道 k 的方差。如图2所示,缩影模型识别来自重叠区域的遥感图像小块,并且遥感图像中不同的潜在变量 s 通过共享参数来压缩这种参数化,其中组件的索引 s 也是参数共享的。因此,缩影是参数 μ, σ 的大网格,混合分量 $s = (s_i, s_j)$ 的参数从位置 s_i, s_j 开始,并根据遥感图像块的大小向左和向下延伸。此外,后验 $p(s|x) \propto p(x|s) p(s)$ 和其他推断是使用卷积/相关来有效执行。

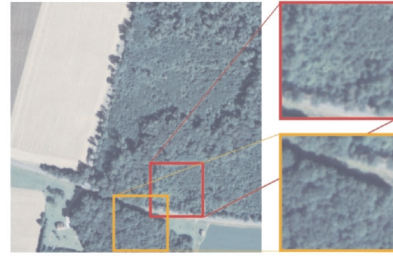


图2 缩影中2个重叠的窗口共享参数

Fig.2 Two overlapping windows in a miniature share parameters

缩影是训练的遥感图像中自相似性的总结。因此,它们应该包含比训练图像少得多的像素数量,但比用于训练它们的缩影图像块要大得多。缩影中的每个像素都包含在大小为 $K \times K$ 的图像块中,并且可以通过遥感图像块映射追溯到许多遥感图像中的不同位置。这种图像映射使得在遥感图像 x 的缩影经过训练后能够将其标签嵌入到缩影中,再将具有标签的缩影组合到遥感图像中,完成基于缩影模型的语义分割。缩影 m, n 中的每个位置都将具有(软)标签指示符 $z_{m,n,l}$,计算公式为:

$$z_{m,n,l} = \sum_t \sum_{i,j:(m,n) \in W_{t,s_j}} p(s_i, s_j | x^t) y_{m-s_i, n-s_j, l}^t \quad (4)$$

其中, W_{s_i, s_j} 表示从 s_i, s_j 开始的缩影窗口,即缩影中属于混合分量 (s_i, s_j) 的 $K \times K$ 坐标集合。从后验中能得出遥感图像块 x^t 与缩影中目标位置 m, n 重叠的每个分量 s 的映射强度。标签 y^t 在块中的对应位

置是 $(m - s_i, n - s_j)$,因此, $y'_{m-s_i, n-s_j, l}$ 被添加到位置 (m, n) 处的标签 l 的计数器 $z_{m, n, l}$ 上。

2) 标签超分辨率。对于1 m分辨率的遥感图像, NLCD的30 m低分辨率土地覆盖产品相当于每个30像素 \times 30像素块被标记为16个类别中的1个, 标签超分辨率就是将每个类别由4个不同的高分辨率标签代替, 形成高分辨率土地覆盖制图。

标签超分辨率技术假设先验知道低分辨率类 c 中高分辨率标签 l 的组合 $p(l|c)$, 下面给出了NLCD类别和4个简化高分辨率标签之间的近似对应关系, 如表1所示。按照低分辨率类中高分辨率标签的统计模型, 可以开发统计标签超分辨率(label super resolution, LSR)推断技术。遥感图像数据 x 由潜在索引 s 的混合索引来建模。使用该索引也对标签 l 和类 c 的联合分布进行建模, 由此可以确定给定低分辨率类别上的已知高分辨率标签分布满足 $p(l|c) = \sum_s p(l|s) p(s|c)$ 。因此, 本研究通过最小化已知的分布 $p(l|c)$ 与模型预测的分布 $p'(l|c) = \sum_s p(l|s) p(s|c)$ 之间的相对熵(KL距离), 通过求解2个概率分布的差异情况得到标签嵌入 $p(l|s)$:

$$p(l|s) = \operatorname{argmax}_{p(l|s)} \sum_c p(c) \sum_l p(l|c) \log p(l|c) \quad (5)$$

其中, $p(c)$ 是数据中观察到的低分辨率类的比例, $p(s|c)$ 作为低分辨率标签 c 数据在 s 上的后验值。首先, 推导了一个EM算法, 用于解决等式(5)中的问题, 该算法使用辅助分布 $q_{l,c}(s)$ 来对函数 $\log \sum_s p(l|s) p(s|c)$ 进行重复定界, 并重新估计了 $p(l|s)$ 的值。为了推导E步骤, 通过观察可以发现:

$$\begin{aligned} \log \sum_s p(l|s) p(s|c) &= \log \sum_s q_{l,c}(s) \frac{p(l|s) p(s|c)}{q_{l,c}(s)} \\ &\geq \log \sum_s q_{l,c}(s) \log \frac{p(l|s) p(s|c)}{q_{l,c}(s)} \end{aligned}$$

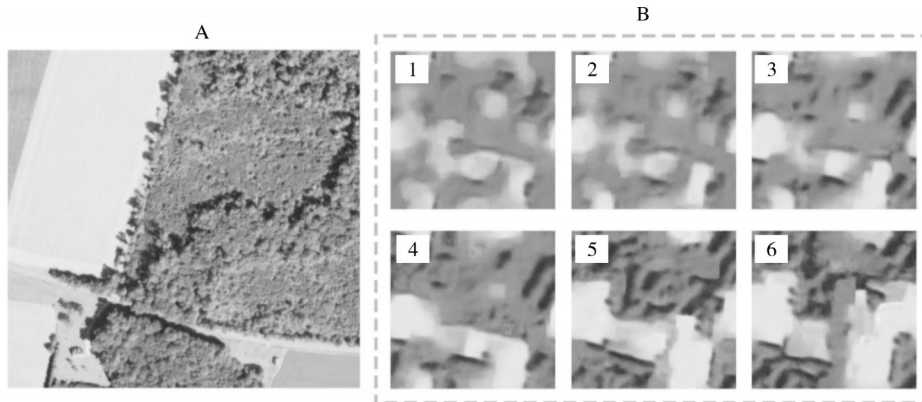
该限定对于所有的分布 $q_{l,c}$ 都是成立的, 并且对以下情况是严格的:

$$q_{l,c}(s) \propto p(l|s) p(s|c) \quad (6)$$

对 $p(l|s)$ 进行优化, 可以得到:

$$p(l|s) \propto p(c) p(l|c) q_{l,c}(s) \quad (7)$$

通过迭代等式(6)和(7), 在 $q_{l,c}(s)$ 和 $p(l|s)$ 上的坐标上收敛到优化标准的局部最大值。如图3所示, 可以看出缩影模型挖掘图像自似性的迭代过程。



A: 512 像素 \times 512 像素的 NAIP 图像 NAIP image of 512 \times 512; B: 1~6 表示 299 \times 299 缩影参数训练的逐步迭代过程 1-6 indicates the step-by-step iterative process of 299 \times 299 microfilm parameter training.

图3 缩影的区域迭代

Fig.3 The area iteration of the miniature

因此, 标签超分辨率所需要的只是概率分布 $p(s|c)$, 该分布能显示出来每个混合成分 s 在低分辨率类别 c 中出现的频率。给出低分辨率的标记数据, 即成对 (x', c') 和经过训练的遥感图像块 x' 的混合模型, 得出的结果是:

$$p(s|c) \propto \sum_{l: c'=c} p(s|x') \quad (8)$$

即该算法遍历了所有的遥感图像块, 查看它们

分配给原型索引 s 的后置, 来计算每个原型索引与每个类的关联次数。

具有参数共享的缩影(epitomics)表示法在此还有额外的优势。对于标准的高斯混合的图像块(patches), 可以完成的标签超分辨率的水平是由分析中使用的遥感图像块缩影 x 的大小定义的, 因为所有推理都是在遥感图像块索引 s 的水平上进行的, 而不是在单个像素上。因此, 为了获得在单个像素水

平的标签超分辨率,可以通过将统计数据 $p(m, n|c)$ 分配给缩影中的各个位置,而不是使用整个缩影图像块的统计信息。

$$p(m, n|c) \propto \sum_i \sum_{i,j} p(m-i, n-j|x^i)[c^i=c] \quad (9)$$

其中, $p(\cdot, \cdot|x^i)$ 是后验位置,式(9)表示计算缩影中的每个像素中低分辨率标签 c 在遥感图像块中映射的次数。考虑到更大空间上下文中的所有遥感图像块,缩影中各个像素可以获得与其邻居不同的统计数据,这允许为整个缩影推断出高分辨率标签。

综上所述,LSR算法使用 $K \times K$ 图像块的缩影模块,使用公式(8)在单个像素级别嵌入类别标签,然后使用迭代公式(6)和(7)的EM算法求解不同分辨率标签 $Z_{s_i, s_j, l} = p(l|s_i, s_j)$ 的可能性,使用公式(9)代替 $p(s|c)$ 。一旦为缩影中的每个位置计算了高分辨率标签的估计值,就可以使用公式(2)预测图像中的标签。此过程对标记有低分辨率类的图像中重复模式的频率执行概率推理,以对这些模式中的单个像素进行推理。最后实现在每个像素上进行标签超分辨率,生成高分辨率土地覆盖分类标签(表1)。

表1 从低分辨率NLCD类别标签到高分辨率类别标签的映射统计数据

Table 1 Mapping statistics from low-resolution NLCD class labels to high-resolution class labels

NLCD类别名称 NLCD class name	类别标签 Category label	近似类别频率/% Approximate class frequencies			
		水 Water	树冠类 Tree canopy	低植被类 Low vegetation	不透水面类 Impervious surface
开阔水域 Open water	水 Water	98	2	0	0
城市开放空间 Developed, open space	复合型 Mixed	0	39	49	12
低密度城市 Developed, low intensity	复合型 Mixed	0	31	34	35
中密度城市 Developed, medium intensity	不透水面 Impervious surface	1	13	22	64
高密度城市 Developed, high intensity	不透水面 Impervious surface	0	3	7	90
荒地 Barren land	低矮植被 Low vegetation	5	13	43	40
落叶林 Deciduous forest	树冠 Tree canopy	0	93	5	0
常绿林 Evergreen forest	树冠 Tree canopy	0	95	4	0
混交林 Mixed forest	树冠 Tree canopy	0	92	7	0
灌木/灌丛 Shrub/scrub	树冠 Tree canopy	0	58	38	4
草地/草本植被 Grassland/herbaceous	低矮植被 Low vegetation	1	23	54	22
牧场/枯草 Pasture/hay	低矮植被 Low vegetation	0	12	83	3
栽培作物 Cultivated crops	低矮植被 Low vegetation	0	5	92	1
森林沼泽 Woody wetlands	树冠 Tree canopy	0	94	5	0
挺水湿地 Emergent wetlands	树冠 Tree canopy	8	86	5	0

1.4 土地覆盖分类图后处理

在标签超分辨率的最终结果输出之后,拟合一个小型神经网络作为后处理步骤。如图4所示,该网络是1个具有5个64通道的 3×3 卷积层,激活层是Relu函数的完全卷积网络(fully convolutional network, FCN)和1个逻辑回归分类器,用来提高模型的预测结果^[22]。向经过训练的网络中输入标签超分辨率的预测结果和NAIP图像,再对相同的图像进行预测,从而得出最终的土地覆盖估计值。由于这种模型具有小的感受野(11×11)、仅对局部纹理敏感,所以不会完全适应缩影模型以及标签超分辨率算法的输出,导致该模型对标签超分辨率方法产生的均匀区域内的斑点噪声和低分辨率类块之间的边界等误差^[26]不敏感。

1.5 多时相土地覆盖制图变化检测方法

语义变化检测是一种多分类问题,通过对比不同时相土地覆盖分类图像之间的差异得出变化检测结果。变化检测的预测图(predictions, P)的表达式为:

$$P = N \times T_1 + T_2 \quad (10)$$

其中, T_1, T_2 代表前后2个不同时相的土地分类标签图像, N 表示高分辨率土地分类标签类别, $N=4$ 。

经历过变化的单个像素既会在一个时间点的数据中“获得”1个类,又会在另1个时间点数据中“失去”1个类。因此,最后会生成2张独立的土地覆盖的预测图。土地覆盖变化分为减少的土地覆盖类别图和增加的土地覆盖类别图,减少的土地覆盖图用公式(11)表示,增加的土地覆盖类别用公式(12)表示。

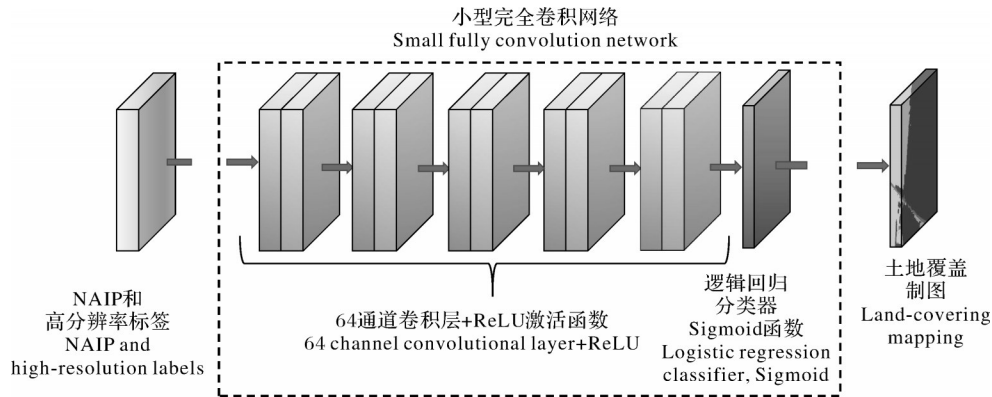


图4 后处理完全卷积网络结构

Fig.4 Post-processing fully convolutional network structure

$$\text{loss} = \frac{P}{4} \quad (11)$$

$$\text{gain} = \frac{P}{25} \quad (12)$$

1.6 模型评价

为了评价标签超分辨率算法对于多时相语义变化检测的性能,采用像素准确率(pixel accuracy, P_A)和平均交并比(mean intersection over union, mIoU)精度指标衡量算法检测结果的好坏。其中, P_A 为预测正确类别的像素点与总像素点的比值,计算公式为:

$$P_A = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (13)$$

为了对变化检测进行评估,确定2个NAIP图像之间的8种高分辨率的变化类型:(1)水类的增加和损失;(2)树冠类的增加和损失;(3)低矮植被的增加和减少;(4)不透水面的增益和损失。评分指标是8个类别的预测标签和地面真实标签之间的交并比(IoU),即4个目标类别中每个的损失和增益,不包括没有变化的类别。虽然没有计算无变化类的IoU,但对没有变化的像素的变化预测仍然会降低公式(14)中的分数。其中,IoU表示实际类别样本和预测类别样本的交集和并集之比,计算公式为:

$$\text{IoU} = \frac{T_P}{T_P + F_P + F_N} \quad (14)$$

将8个类别的IoU累加之后进行平均处理得到评价指标平均交并比(mean intersection over union, mIoU):

$$\text{mIoU} = \frac{1}{k} \sum_{i=1}^k \text{IoU} \quad (15)$$

公式(13)~(15)中: T_P 为正确分类的正例; T_N 为正确分类的负例; F_P 为错分成正例的负例; F_N 为错分成负例的正例; k 为土地覆盖的总类别数。

1.7 模型环境和参数

试验环境为NVIDIA Tesla V100显卡,16 G显存, GPU加速库采用CUDA-10.1。在Ubuntu16.04.7下另外安装Anaconda3并创建既有Python 3.6和Pytorch 1.5.1的软件环境。使用Pytorch框架搭建整个网络,输入的图像大小均为512像素×512像素,在将图像送入网络之前,低分辨率输入增强和随机颜色增强将被应用于传入的遥感图像。构建了1个299像素×299像素的缩影模型。试验中的具体参数见表2。为了多元化训练,将缩影模型优化准则从所有数据的对数似然改变为每批最差子集(试验中最低的25%中的数据集)的对数似然。将标签超分辨率土地分类结果图传入到后处理全卷积神经网络模型中。土地覆盖分类图后处理中的FCN模型的具体参数如表2所示。

2 结果与分析

2.1 标签超分辨率方法比较

表3所示为采用不同模块处理对模型性能的影响。其中,Epitomic LSR表示本研究所提到的基于缩影的标签超分辨率LSR算法,Color表示输入颜色增强,LR表示低分辨率输入增强,NS表示神经平滑网络。“Epitomic LSR+LR+Color”表示在数据输入模型之前进行数据增强操作,“Epitomic LSR+LR+NS”表示在缩影模型输出后进行神经平滑操作,“Epitomic LSR+LR+Color+NS”表示将数据增强和神经平滑网络联合进行训练。

从表3可以看出,不同模块的添加提高了模型的性能。具体而言,对于本研究所提出的标签超分辨率模型,低分辨率数据增强、颜色增强以及在模型后面添加的神经平滑网络都提高了模型在数据集上面

表2 FCN模型的具体参数

Table 2 Specific parameters of the FCN model

模型 Model	参数名称 Parameter name	参数值 Parameter	
缩影模型 Epitomes model (299×299)	可变大小的块 w Variable size block w	11×11~31×31	
	平滑 $p(x^i s)$ 概率依据 Smoothed $p(x^i s)$ Probability basis	$(w/11)^2$	
	均值(μ) Mean	0.5+unif(0, 0.1)	
	方差($1/\sigma^2$) Variance	初始值 Initial value	10
		范围 Range	1~100
	先验在对数域参数化 Prior	初始值 Initial value	0
		范围 Range	-4~4
	学习率 Learning rate	0.003	
	优化器 Optimizer	Adam	
	迭代次数 Number of iterations	50 000	
批处理大小 Batch size	256		
FCN模型 FCN model	位置推广阈值 Location promotion threshold	10^{-8}	
	计数器重置阈值(δ) Counter reset threshold	0.05	
	训练 Epoch	50	
	批处理大小 Batch size	64	
	学习率(α) Learning rate(α)	0.001	

表3 不同模块处理对模型性能的影响

Table 3 Effects of different module processing on model performance

方法 Method	像素准确率 P_A	平均交并比 mIoU
Epitomic LSR	85.4	58.6
Epitomic LSR+LR+Color	86.3	63.5
Epitomic LSR+LR+NS	86.8	65.2
Epitomic LSR+LR+Color+NS	87.6	68.1

的像素准确率(P_A)和平均交并比(mIoU)。当仅使用标签超分辨率缩影 LSR 算法, P_A 为 85.4%, mIoU 为 58.6%。由于低分辨率数据增强可以用来平滑高分辨率图像的局部噪声源, 所以当引入数据增强以及输入图像随机颜色增强后, 标签超分辨率的 P_A 和 mIoU 分别提高到 86.3%、63.5%。后处理是一个小型神经网络, 具有小的感受野并且仅对局部纹理敏感, 这会导致对缩影标签超分辨率算法产生的区域内斑点噪声和低分辨率类块之间的边界错误不敏感, 从而提升标签超分辨率的结果, P_A 和 mIoU 分别提高到 86.8%、65.2%。将这 2 种方法同时加入模型中, 模型精准度达到最好的效果, P_A 和 mIoU 分别为 87.6%、68.1%。

2.2 与其他多时相语义变化检测方法的比较

将本研究提出的算法与 NLCD diff、FCN/逐块 (FCN/tile)、FCN/整体 (FCN/all)、U-Net/逐块 (U-Net/tile)、U-Net/整体 (U-Net/all) 等 5 种基线多时相语义变化检测方法进行比较。依据 4 个目标类别的增益和损失的 IoU 得分与最终的 mIoU 进行分析, 以验证所提出的多时相语义变化检测算法的有效性。表 4 为本研究提出的方法在交并比 (IoU) 得分, 其中, -W 和 +W 分别表示水类的损失和增益的; +TC 和 -TC 别表示树冠类的损失和增益; -LV 和 +LV 分别表示低植被类的损失和增益; -I 和 +I 分别表示不透水表面类的损失和增益。

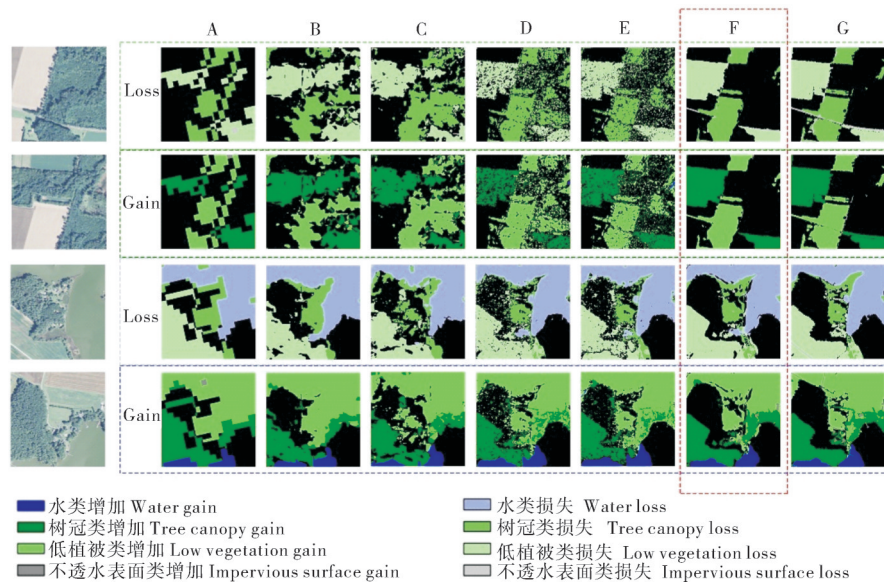
由表 4 可知, 本研究提出的基于缩影表示的标签超分辨率模型 mIoU 为 59.2%, 优于其他基线方法。NLCD diff 方法由于只给出粗略的变化检测, 许多变化都发生在小而薄的地方, 不能很好地预测土地覆盖变化。U-Net 具有较大的感受野, 能进行模糊预测, 但表现效果低于具有小的感受野的 FCN。本研究提出的方法正是基于此原因, 放弃远距离关系, 直接建模小图像块。由于缩影看起来类似于被训练的图像块, 大大增加了模型的可解释性, 且该图像分割生成公式允许与其他统计程序相结合, 在有或没有高分辨率监督的情况下推断潜在变量空间中的标签, 这些使得在标签超分辨率设置方面的性能优于卷积神经网络 (convolutional neural networks, CNNs)。可以观察到数据增强比没有进行数据增强处理的所有模型效果更好, 表明数据增强可以解决高分辨率图像中局部噪声影响模型准确度的问题。最后再用小的 FCN 网络进行平滑, 提供了所需丰富而重要的浅层信息, 展示出良好的性能。如图 5 所示, 不透水面很难达到高分, 大多数错误预测来自于和其他类之间的模糊, 这可能是由于原始 NLCD 中大量的不匹配边界造成的。

3 讨论

本研究中提出了一种基于缩影的多时相遥感语义变化检测方法。缩影模型采用简单的图像块混合模型, 类似被训练的图像, 因此该模型是可解释的。基于该优点, 建模所需要的不变性可以直接用额外的隐藏变量即通过向该模型中加入直观且易于理解的先验知识来实现^[27]。由于缩影模型从像素中学习遥感图像特征, 可以有效地使用小图像块来执行视觉任务。通过基于缩影表示的模型进行弱监督语义

表4 几种多时相语义变化检测方法检测结果比较

方法 Method	-W	-TC	-LV	-I	+W	+TC	+LV	+I	mIoU
NLCD diff	14.8	16.7	28.2	1.4	3.1	0.1	10.6	36.2	13.9
FCN/tile	64.1	43.6	40.7	9.1	25.5	7.3	30.2	52.8	34.2
FCN/all	58.4	64.4	60.1	40.0	29.2	18.1	60.7	71.6	50.3
U-Net/tile	27.5	48.3	47.7	29.1	15.8	14.8	35.1	51.6	33.7
U-Net/all	32.5	48.4	47.6	30.7	23.7	20.5	34.2	51.7	36.1
Epitomic LSR	71.3	50.0	72.5	26.6	58.2	68.7	61.4	20.9	53.7
Epitomic LSR+LR+Color	74.3	52.2	70.1	37.2	62.6	67.3	63.0	22.1	56.1
Epitomic LSR+LR+NS	77.2	58.3	75.2	32.7	59.2	72.1	64.2	19.4	57.3
Epitomic LSR+LR+Color+NS	78.0	62.7	76.3	29.1	62.9	68.8	71.5	24.6	59.2



A~E:NLCD diff,FCN/逐块,FCN/整体,U-Net/逐块,U-Net/整体的检测结果 Detection of NLCD diff,FCN/tile,FCN/all,U-Net/tile,U-Net/all;F:本研究的结果 Results of the method proposed herein;G:真实的地面标签 The true ground label.

图5 几种多时相语义变化检测方法检测结果示例

Fig.5 Examples of detection results from several multi temporal semantic change detection methods

分割,在土地分割结果上应用标签超分辨率算法(label super resolution,LSR)生成高分辨率土地覆盖地图。再通过拟合1个小型FCN网络来平滑生成的高分辨率遥感图像标签中的噪声和误差,最后比较2个不同时间点的土地覆盖制图生成结果变化图进而实现多时相语义变化检测。基于缩影的多时相遥感语义变化检测方法实现了对2个时相的遥感影像变化区域的自动提取,可以得到变化区域具体类型的语义信息。试验结果表明,该方法在像素准确率(P_A)和平均交并比(mIoU)的评价指标上表现良好。

未来遥感应用的重要基础是实时的大数据服务平台。为提高本研究方法的实用性,可选择利用实

时遥感数据平台结合基于缩影的多时相语义变化检测技术实现实时观测和开展区域或土地覆盖变化检测的研究。GEE平台公共数据库中提供了目前经常使用的遥感影像数据、地理空间数据集,可以快速、批量处理海量的遥感大数据。使用GEE平台进行长时间序列和大范围土地利用变化遥感监测研究和应用,具有很强的平台优势和技术先进性。

参考文献 References

- [1] 阎守邕.现代遥感技术系统及其发展趋势[J].环境遥感,1995(1):52-62.YAN S Y.Modern remote sensing technical system and its development trends[J].National remote sensing bulletin, 1995(1):52-62(in Chinese with English abstract).

- [2] 张良培, 武辰. 多时相遥感影像变化检测的现状与展望[J]. 测绘学报, 2017, 46(10): 1447-1459. ZHANG L P, WU C. Advance and future development of change detection for multi-temporal remote sensing imagery[J]. Acta geodaetica et cartographica sinica, 2017, 46(10): 1447-1459 (in Chinese with English abstract).
- [3] SOHL T L. Change analysis in the United Arab Emirates: an investigation of techniques [J]. Photogrammetric engineering and remote sensing, 1999, 65(4): 475-484.
- [4] BERBEROGLU S, AKIN A. Assessing different remote sensing techniques to detect land use/cover changes in the eastern Mediterranean [J]. International journal of applied earth observation and geoinformation, 2009, 11(1): 46-53.
- [5] 程腾辉, 范文义, 吴国明. 多时相 PolInSAR 数据土地覆盖类型分类研究[J]. 森林工程, 2022, 38(5): 10-17. CHENG T H, FAN W Y, WU G M. Research on land cover type classification of multi-temporal PolInSAR data [J]. Forest engineering, 2022, 38(5): 10-17 (in Chinese with English abstract).
- [6] MOU L C, BRUZZONE L, ZHU X X. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery [J]. IEEE transactions on geoscience and remote sensing, 2018, 57(2): 924-935.
- [7] ZHANG C X, YUE P, TAPETE D, et al. A multi-level context-guided classification method with object-based convolutional neural network for land cover classification using very high resolution remote sensing images [J/OL]. International journal of applied earth observation and geoinformation, 2020, 88: 102086 [2022-08-02]. <https://doi.org/10.1016/j.jag.2020.102086>.
- [8] YANG K P, XIA G S, LIU Z C, et al. Asymmetric Siamese networks for semantic change detection in aerial images [J]. IEEE transactions on geoscience and remote sensing, 2022, 60: 1-18.
- [9] 吴伟斌, 唐婷, 刘强, 等. 基于语义分割的山地果园道路识别技术研究[J]. 华中农业大学学报, 2022, 41(1): 246-254. WU W B, TANG T, LIU Q, et al. Semantic segmentation based road recognition technology of hilly fruit and tea garden [J]. Journal of Huazhong Agricultural University, 2022, 41(1): 246-254 (in Chinese with English abstract).
- [10] 郭海涛, 卢俊, 袁洲, 等. Siam-DeepLabv3+网络遥感影像语义变化检测方法[J]. 测绘科学技术学报, 2021, 38(6): 597-603. GUO H T, LU J, YUAN Z, et al. A semantic change detection method for remote sensing images based on Siam-DeepLabv3+ [J]. Journal of geomatics science and technology, 2021, 38(6): 597-603 (in Chinese with English abstract).
- [11] 王利民, 刘佳, 季富华. 中国农业遥感技术应用现状及发展趋势[J]. 中国农学通报, 2021, 37(25): 138-143. WANG L M, LIU J, JI F H. Status quo and development trend of agriculture remote sensing technology application in China [J]. Chinese agricultural science bulletin, 2021, 37(25): 138-143 (in Chinese with English abstract).
- [12] 史舟, 梁宗正, 杨媛媛, 等. 农业遥感研究现状与展望[J]. 农业机械学报, 2015, 46(2): 247-260. SHI Z, LIANG Z Z, YANG Y Y, et al. Status and prospect of agricultural remote sensing [J]. Transactions of the CSAM, 2015, 46(2): 247-260 (in Chinese with English abstract).
- [13] CHI M M, PLAZA A, BENEDIKTSSON J A, et al. Big data for remote sensing: challenges and opportunities [J]. Proceedings of the IEEE, 2016, 104(11): 2207-2219.
- [14] MALKIN N, ROBINSON C, JOJIC N. High-resolution land cover change from low-resolution labels: simple baselines for the 2021 IEEE GRSS Data Fusion Contest [DB/OL]. arXiv, 2019: 2101.01154 [2022-08-02]. <https://arxiv.org/abs/2101.01154>.
- [15] YOKOYA N, GHAMISI P, HANSCH R, et al. 2021 data fusion contest: geospatial artificial intelligence for social good technical committees [J]. Geoscience and remote sensing, 2021, 9(4): 274-277.
- [16] FU K, LU W X, DIAO W H, et al. WSF-NET: weakly supervised feature-fusion network for binary segmentation in remote sensing image [J/OL]. Remote sensing, 2018, 10(12): 1970 [2022-08-02]. <https://doi.org/10.3390/rs10121970>.
- [17] WANG S, CHEN W, XIE S M, et al. Weakly supervised deep learning for segmentation of remote sensing imagery [J/OL]. Remote sensing, 2020, 12(2): 207 [2022-08-02]. <https://doi.org/10.3390/rs1202207>.
- [18] PATHAK D, KRÄHENBÜHL P, DARRELL T. Constrained convolutional neural networks for weakly supervised segmentation [C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile, USA. [S.l.]: IEEE, 2016: 1796-1804.
- [19] ROBINSON C, HOU L, MALKIN K, et al. Large scale high-resolution land cover mapping with multi-resolution data [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. [S.l.]: IEEE, 2020: 12718-12727.
- [20] MALKIN K, ROBINSON C, HOU L, et al. Label super-resolution networks [C]//Proceedings of the International Conference on Learning Representations, Apr 30-May 3, 2018, Vancouver, Canada. Vancouver: ICLR, 2018.
- [21] ZHAO M Z, HOU L, LE H, et al. Label super resolution with inter-instance loss [DB/OL]. arXiv, 2019: 1904.04429 [2022-08-02]. <https://arxiv.org/abs/1904.04429>.
- [22] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(4): 640-651.
- [23] JOJIC N, FREY B J, KANNAN A. Epitomic analysis of appearance and shape [C]//Proceedings of the Ninth IEEE International Conference on Computer Vision, October 13-16, 2003, Nice, France. [S.l.]: IEEE, 2008: 34-41.
- [24] MALKIN N, ORTIZ A, JOJIC N. Mining self-similarity: label super-resolution with epitomic representations [C]//Computer

- Vision-ECCV 2020: 16th European Conference, August 23-28, 2020, Glasgow, UK. New York: ACM, 2020: 531-547.
- [25] HOMER C, DEWITZ J, YANG L M, et al. Completion of the 2011 national land cover database for the conterminous United States: representing a decade of land cover change information [J]. *Photogrammetric engineering & remote sensing*, 2015, 81(5): 345-354.
- [26] ROBINSON C, MALKIN K, HU L, et al. Weakly supervised semantic segmentation in the 2020 IEEE GRSS data fusion contest[C]//IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, September 26-October 2, 2020, Waikoloa, HI, USA. Waikoloa: IEEE, 2021: 7046-7049.
- [27] FREY B J, JOJIC N. Transformed component analysis: joint estimation of spatial transformations and image components[C]//Proceedings of the Seventh IEEE International Conference on Computer Vision, September 20-27, 1999, Kerkyra, Greece. Kerkyra: IEEE, 2002: 1190-1196.

A method of detecting multitemporal semantic changes based on epitomes

JING Weipeng¹, WANG Jian¹, ZHANG Wenjun¹, GU Juntao², CHEN Guangsheng¹

1. *College of Information and Computer Engineering, Northeast University of Forestry, Harbin 150040, China;*

2. *Heilongjiang Research Center for Cyberspace, Harbin 150090, China*

Abstract The detection of multitemporal semantic changes is often used to monitor changes in agricultural ecology and to track the development of agricultural land because it uses semantic information to analyze the specific types of changes. A method of detecting multitemporal semantic changes using weak labels with noise and low resolution instead of high-resolution labels was proposed to solve the problem that the scarcity of high-resolution remote sensing image labels and the slow growth of labeling technology limit the development of detecting multitemporal semantic changes. First, low resolution satellite data were used to smooth the quality differences of high-resolution remote sensing image inputs. Secondly, the high-resolution remote sensing image classification map was estimated by combining the epitomes model and the label super-resolution algorithm as a statistical inference algorithm, and a small FCN network was fitted to post-process the remote sensing image classification map generated to improve its classification. Finally, the results of detecting change were obtained by comparing the differences between different simultaneous land cover classification images. The results showed that the proposed method improved the mean intersection over union (mIoU) by 8.9 percentage points compared with other methods of detecting multitemporal semantic changes, and detected the changes of land cover classification effectively.

Keywords weak supervision; label super-resolution; epitomes; map of land cover changes; semantic segmentation; remote sensing image; detecting changes

(责任编辑:陆文昌)