



(12) 发明专利

(10) 授权公告号 CN 116524346 B

(45) 授权公告日 2024. 08. 09

(21) 申请号 202310203834.1

G06V 20/70 (2022.01)

(22) 申请日 2023.03.06

G06V 10/764 (2022.01)

(65) 同一申请的已公布的文献号

G06N 3/0464 (2023.01)

申请公布号 CN 116524346 A

G06N 3/047 (2023.01)

(43) 申请公布日 2023.08.01

G06N 3/09 (2023.01)

G06N 3/084 (2023.01)

(73) 专利权人 西北工业大学

(56) 对比文件

地址 710072 陕西省西安市友谊西路

CN 111582043 A, 2020.08.25

(72) 发明人 张艳宁 张秀伟 杨一哲 于雷

CN 112308860 A, 2021.02.02

田牧 安博远 邢颖慧

审查员 周勇攀

(74) 专利代理机构 西安凯多思知识产权代理事

务所(普通合伙) 61290

专利代理师 刘新琼

(51) Int. Cl.

G06V 20/10 (2022.01)

G06V 10/82 (2022.01)

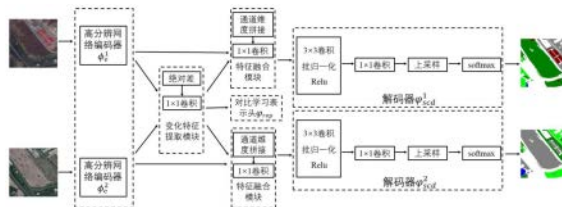
权利要求书3页 说明书10页 附图2页

(54) 发明名称

基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法

(57) 摘要

本发明涉及一种基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法;构建了基于高分辨网络的简单可扩展的直接语义变化检测模型来进行语义变化检测,并在变化检测上对比学习损失作为监督,有效挖掘由于类别不平衡而导致的难分类样本,提高网络对难分类样本的分类性能;在对比学习中,采用半难半易的采样策略,使网络在关注难分类样本的同时易于收敛。通过半难半易的采样策略,对比损失可以指导网络为变化检测提供适当的监督,特别是更加关注难以正确分成变化类或未变化类的样本,从而提高模型的整体语义变化性能。本发明在变化区域上细节更加精确,形状更加完整,同时对语义变化类别的分类更加准确。



1. 一种基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法,其特征在于步骤如下:

步骤1:构建基于高分辨网络的简单可扩展的直接语义变化检测模型来进行语义变化检测;

将 T^1 和 T^2 时相的遥感图像 I^1 和 I^2 输入两个共享权重的高分辨网络编码器 ϕ_e^1 和 ϕ_e^2 得到语义上下文特征对 f_{sem}^1 和 f_{sem}^2 ,将 f_{sem}^1 和 f_{sem}^2 输入变化特征提取模块得到变化特征 f_{cd} ;对于 T^1 时相,将 f_{sem}^1 和 f_{cd} 输入特征融合模块,先对 f_{sem}^1 和 f_{cd} 作绝对差得到 d_{cd} ,再将 d_{cd} 进行1次 1×1 卷积操作得到语义变化特征 f_{scd}^1 ;对于 T^2 时相,将 f_{sem}^2 和 f_{cd} 输入特征融合模块得到语义变化特征 f_{scd}^2 ;最后将 f_{scd}^1 和 f_{scd}^2 输入两个解码器 ϕ_{scd}^1 和 ϕ_{scd}^2 得到 T^1 和 T^2 时相的语义变化检测结果 O_{scd}^1 和 O_{scd}^2 ;

步骤2:使用对比学习损失监督变化检测;

在变化特征提取模块后加入对比学习特征表示头 Φ_{rep} ,将变化特征 f_{cd} 输入对比学习特征表示头 Φ_{rep} 得到密集变化特征表示 Z_{rep} , Φ_{rep} 包含2个卷积层;第一个卷积层包含1个步幅为1,填充为2的 3×3 卷积,该卷积操作将输入变化特征 f_{cd} 通道数变为原来1/4,分辨率保持不变,即 $H_0 \times W_0 \times 150$,以及1个批归一化操作和1修正线性单元;第二个卷积层包含1个步幅为1的 1×1 卷积操作,将输入特征向量通道数变为256,分辨率保持不变,即 $H_0 \times W_0 \times 256$;由于对比损失仅在训练期间提供监督,因此对比学习特征表示头在推理期间被移除;用半难半易的采样策略在变化特征表示 Z 中对变化类别和未变化类别的样本对应的特征向量进行采样以计算对比学习损失函数 L_c :

$$L_c = -\frac{1}{C \times A} \sum_{c=0}^{C-1} \sum_{i=1}^A \log \left[\frac{e^{\langle z_{ca}, z_{ca}^+ \rangle / \tau}}{e^{\langle z_{ca}, z_{ca}^+ \rangle / \tau} + \sum_{b=1}^B e^{\langle z_{ca}, z_{ca,b}^- \rangle / \tau}} \right]$$

其中,由于类别为变化类别和未变化类别,所以 $\|C\|=2$; z_{ca} 是第 c 类第 a 个锚点的特征向量, z_{ca}^+ 是第 c 类的正样本,是所有第 c 类样本特征向量的平均值, $z_{ca,b}^-$ 是第 c 类第 a 个锚点的第 b 个属于其他类别的负样本的特征向量;对于 C 中的每个类别,锚点和锚点的负样本对应的特征向量都采集自变化特征表示 Z_{rep} ,每个类别都有 A 个锚点,并且每个锚点都有一个正样本和 B 个负样本; $\langle \cdot, \cdot \rangle$ 是两个特征向量之间的余弦相似度,用于度量两个特征向量之间的距离,范围为-1到1, $\tau=0.5$ 是温度系数;对于当前训练批量中的每个类别,通过优化网络参数使得 L_c 最小,来拉近该类别的锚点与其正样本的距离,而推远该类别的锚点与其负样本的距离;

步骤3:通过使总体的损失函数 L 最小来优化模型的参数,总体的损失函数 L 由 T^1 和 T^2 的语义变化检测损失 L_{scd}^1 和 L_{scd}^2 以及变化检测上的对比学习损失函数 L_c 构成;语义变化检测损失 L_{scd}^1 和 L_{scd}^2 为交叉熵损失,描述为:

$$L_{scd}^1 = -\frac{1}{T} \sum_{t=1}^T y_t^1 \log(p_t^1)$$

$$L_{scd}^2 = -\frac{1}{T} \sum_{t=1}^T y_t^2 \log(p_t^2)$$

其中T为像素的个数, y_t^1 和 p_t^1 , y_t^2 和 p_t^2 分别表示真值标签和解码器 φ_{scd}^1 和 φ_{scd}^2 在第t个像素上的预测概率;

总体的损失函数L描述为:

$$L = L_{scd}^1 + L_{scd}^2 + L_c$$

得到总体的损失函数后进行反向传播,使用AdamW优化器进行优化,重复迭代直至迭代次数达到设置初始值时判定完成训练。

2. 根据权利要求1所述的基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法,其特征在于:步骤1具体如下:

步骤1-1:对于 T^1 时相,遥感图像 I^1 通过卷积模块 M_0 得到的维度为 $H_0 \times W_0 \times 64$ 的特征图 m_0^1 , $H_0 = H_{input}/4$ 与 $W_0 = W_{input}/4$,其中 H_{input} 与 W_{input} 为输入图像 I^1 的高和宽;

步骤1-2:将 m_0^1 输入高分辨网络编码器 ϕ_e^1 ,高分辨网络编码器 ϕ_e^1 包括4层,第i层包括i个卷积模块 $M_{i,j}$, $i \geq 1, 0 \leq j < i$;卷积模块 $M_{i,j}$ 输出的特征图定义为 $m_{i,j}^1$,特征图 $m_{i,j}^1$ 的分辨率记为 $H_{i,j} \times W_{i,j}$,通道数记为 $C_{i,j}$,其中 $H_{i,j} = H_0/2^j$, $W_{i,j} = W_0/2^j$, $C_{i,j} = 40 \times 2^j$;

当 $i \geq 2$ 时,将第i-1层卷积模块 $M_{i-1,y}$ 的输出 $m_{i-1,y}^1$ 经过如下处理统一输入分辨率及通道数后采用加法融合输入第i层的卷积模块 $M_{i,j}$, $y \in [0, i-2]$;

当 $y < j$ 时,对特征图 $m_{i-1,y}^1$ 进行j-y次步长为2的跨步卷积,每次跨步卷积通过 3×3 卷积使得特征图的通道数翻倍,分辨率减半;

当 $y = j$ 时,对特征图 $m_{i-1,y}^1$ 进行 3×3 卷积进行特征的提取且通道数与分辨率不变;

当 $y > j$ 时,对特征图 $m_{i-1,y}^1$ 进行 3×3 卷积,并将通道数变为 40×2^j ,且使用双线性插值上采样操作使得分辨率变为 $H_0/2^j \times W_0/2^j$;

通过高分辨编码器中所有的卷积模块后,得到的4个不同尺度的特征图 $m_{4,k}^1$, $k = 0, 1, 2, 3$; $m_{4,k}^1$ 的分辨率与通道数分别为 $H_0/2^k \times W_0/2^k$ 和 40×2^k ;将 $m_{4,k}^1$ 通过上采样操作统一分辨率,再沿通道维度进行拼接,得到语义上下文特征对 f_{sem}^1 和 f_{sem}^2 ;

步骤1-3,将语义上下文特征对 f_{sem}^1 和 f_{sem}^2 输入变化特征提取模块,对 f_{sem}^1 和 f_{sem}^2 作绝对差 d_{cd} ,再将 d_{cd} 进行1次 1×1 卷积操作,得到尺度为 $H_0 \times W_0 \times 600$ 变化特征 f_{cd} ;

步骤1-4,对于 T^1 时相,将 f_{sem}^1 和 f_{cd} 输入特征融合模块,特征融合模块包含1个沿通道维度的拼接操作和1个改变通道数的 1×1 卷积操作;首先沿通道维度进行拼接,再通过1次 1×1 卷积操作将通道数压缩为原来的1/2,得到尺度为 $H_0 \times W_0 \times 600$ 的语义变化特征 f_{scd}^1 ;

步骤1-5,将 f_{scd}^1 输入解码器 φ_{scd}^1 ,解码器包含2个卷积层,一个双线性插值上采样操作和一个softmax操作;第一个卷积层包含1个步幅为1,填充为2的 3×3 卷积,该卷积操作将输

入特征向量通道数变为原来1/4,分辨率保持不变,即 $H_0 \times W_0 \times 150$,以及1个批归一化操作和1修正线性单元;第二个卷积层包含1个步幅为1的 1×1 卷积操作,将输入特征向量通道数变为语义变化检测类别数,分辨率保持不变,即 $H_0 \times W_0 \times 7$;双线性插值上采样操作将输入特征向量的分辨率恢复为输入图像的分辨率,即 $H_{input} \times W_{input} \times 7$,最后用softmax操作进行归一化,得到时相的语义变化检测结果 O_{scd}^1 ;T²时相同理得到 O_{scd}^2 。

3. 根据权利要求2所述的基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法,其特征在于:所有的卷积模块都由 3×3 卷积层、批归一化层与修正线性单元组成。

4. 根据权利要求1所述的基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法,其特征在于:步骤2所述的半难半易的采样策略是对每个类别的锚点,都采样一半的难分类样本和一半的易分类样本,对于每个锚点的负样本,都采用一半的难分类样本和一半的易分类样本;用阈值 δ 划分难分类样本和易分类样本;难分类样本在其对应类别上的二值变化检测预测概率小于阈值 δ ,而易分类样本在其对应类别上的二值变化检测预测概率大于阈值 δ ,其中二值变化检测预测概率来自编码器softmax归一化后的结果。

5. 一种计算机系统,其特征在于包括:一个或多个处理器,计算机可读存储介质,用于存储一个或多个程序,其中,当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现权利要求1所述的方法。

6. 一种计算机可读存储介质,其特征在于存储有计算机可执行指令,所述指令在被执行时用于实现权利要求1所述的方法。

基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法

技术领域

[0001] 本发明属于遥感图像处理技术领域,具体涉及一种基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法。

背景技术

[0002] 语义变化检测是遥感图像解译中的一个关键和具有挑战性的任务,它是一种利用同一地理位置多时相遥感图像来检测、定位和分析地球表面土地覆盖类型“从-到”语义变化的方法,该方法在城市规划、环境监测和灾害评估等领域发挥着重要作用。

[0003] 近年来,随着深度学习技术的快速发展和大量多时相高分辨率遥感图像的产生,基于深度学习的语义变化检测方法取得了长足的进步。与传统的语义变化检测方法相比,其语义变化检测性能有了明显的提高。为了同时解决语义分割和变化检测问题,现有的基于深度学习的方法大多侧重于设计合适的网络结构来有效地编码和整合语义上下文信息和变化信息。例如,Dault等人在文献“Multitask learning for large-scale semantic change detection.”中探索并比较了4种常见的基于深度学习的语义变化检测网络结构,分别是:比较双时相语义分割结果得到语义变化检测结果,直接进行语义变化检测,分别进行语义分割和变化检测,分别进行语义分割和变化检测并在解码器阶段将语义分割信息提供给变化检测。Peng等人在文献“Scdnet:Anovel convolutional network for semantic change detection in high resolution optical remote sensing imagery.”中提出了SCDNet,在编码器阶段,其采用共享权重的Resnet34提取语义上下文信息和变化信息,在解码器阶段整合语义上下文信息和变化信息直接进行语义变化检测。Zheng等人在文献“Changemask:Deep multi-task encoder-transformer-decoder architecture for semantic change detection.”中将SCD解耦为两个语义分割任务和一个变化检测任务,然后采用EfficientNet作为编码器,UNet++作为语义分割的解码器,Transformer提取变化信息,Unet++作为变化检测的解码器,最后用语义分割结果和变化检测结果产生语义变化检测结果。

[0004] 但是目前深度学习的语义变化检测方法还存在着一些问题:1.在现有的语义变化检测方法中,一部分方法将语义变化检测任务拆分成语义分割任务和变化检测任务,通过整合语义分割和变化检测的结果得到最终的语义变化检测结果,这种方法会占据大量的内存和计算资源,另一部分方法直接进行语义变化检测,但是不能够在整合语义分割和变化检测信息的同时保持模型的简洁高效和可扩展性,因此设计一个简单可扩展的网络直接进行语义变化检测非常有必要;2.直接进行语义变化检测时,二值变化检测的效果会极大程度上影响整体的语义变化检测的效果,因此在进行语义变化检测时为变化检测加入额外的监督是非常有必要的;3.高分辨率语义变化检测数据集在未变化类别和变化类别上存在严重的类别不平衡问题,具体来说,未变化类别的样本数在总体样本数中占比超过90%,远大于变化类别的样本数。在训练过程中,网络由占比高的未变化类别主导,导致在占比低的变

化类别上性能不佳,因此在训练过程中,会出现大量难以正确进行分类的样本,如何关注和利用这些难样本,让网络能够对它们进行正确地分类是重点和难点。

发明内容

[0005] 要解决的技术问题

[0006] 针对目前语义变化检测结果精度不足的问题,本发明提供一种基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法。

[0007] 技术方案

[0008] 一种基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法,其特征在于步骤如下:

[0009] 步骤1:构建基于高分辨网络的简单可扩展的直接语义变化检测模型来进行语义变化检测;

[0010] 将 T^1 和 T^2 时相的遥感图像 I^1 和 I^2 输入两个共享权重的高分辨率网络编码器 ϕ_e^1 和 ϕ_e^2 得到语义上下文特征对 f_{sem}^1 和 f_{sem}^2 ,将 f_{sem}^1 和 f_{sem}^2 输入变化特征提取模块得到变化特征 f_{cd} ;对于 T^1 时相,将 f_{sem}^1 和 f_{cd} 输入特征融合模块,先对 f_{sem}^1 和 f_{sem}^2 作绝对差得到 d_{cd} ,再将 d_{cd} 进行1次 1×1 卷积操作得到语义变化特征 f_{scd}^1 ;对于 T^2 时相,将 f_{sem}^2 和 f_{cd} 输入特征融合模块得到语义变化特征 f_{scd}^2 ;最后将 f_{scd}^1 和 f_{scd}^2 输入两个解码器 ϕ_{scd}^1 和 ϕ_{scd}^2 得到 T^1 和 T^2 时相的语义变化检测结果 O_{scd}^1 和 O_{scd}^2 ;

[0011] 步骤2:使用对比学习损失监督变化检测;

[0012] 在变化特征提取模块后加入对比学习特征表示头 Φ_{rep} ,将变化特征 f_{cd} 输入对比学习特征表示头 Φ_{rep} 得到密集变化特征表示 Z_{rep} , Φ_{rep} 包含2个卷积层;第一个卷积层包含1个步幅为1,填充为2的 3×3 卷积,该卷积操作将输入变化特征 f_{cd} 通道数变为原来1/4,分辨率保持不变,即 $H_0 \times W_0 \times 150$,以及1个批归一化操作和1修正线性单元;第二个卷积层包含1个步幅为1的 1×1 卷积操作,将输入特征向量通道数变为256,分辨率保持不变,即 $H_0 \times W_0 \times 256$;由于对比损失仅在训练期间提供监督,因此对比学习特征表示头在推理期间被移除;用半难半易的采样策略在变化特征表示 Z 中对变化类别和未变化类别的样本对应的特征向量进行采样以计算对比学习损失函数 L_c :

$$[0013] \quad L_c = -\frac{1}{C \times A} \sum_{c=0}^{C-1} \sum_{i=1}^A \log \left[\frac{e^{(z_{ca}, z_{ca}^+)/\tau}}{e^{(z_{ca}, z_{ca}^+)/\tau} + \sum_{b=1}^B e^{(z_{ca}, z_{ca,b}^-)/\tau}} \right]$$

[0014] 其中,由于类别为变化类别和未变化类别,所以 $\|C\|=2$; z_{ca} 是第 c 类第 a 个锚点的特征向量, z_{ca}^+ 是第 c 类的正样本,是所有第 c 类样本特征向量的平均值, $z_{ca,b}^-$ 是第 c 类第 a 个锚点的第 b 个属于其他类别的负样本的特征向量;对于 C 中的每个类别,锚点和锚点的负样本对应的特征向量都采集自变化特征表示 Z_{rep} ,每个类别都有 A 个锚点,并且每个锚点都有一个正样本和 B 个负样本; $\langle \cdot, \cdot \rangle$ 是两个特征向量之间的余弦相似度,用于度量两个特征向量之间的距离,范围为-1到1, $\tau=0.5$ 是温度系数;对于当前训练批量中的每个类别,通过优化网络参数使得 L_c 最小,来拉近该类别的锚点与其正样本的距离,而推远该类别的锚点与

其负样本的距离；

[0015] 步骤3:通过使总体的损失函数L最小来优化模型的参数,总体的损失函数L由 T^1 和 T^2 的语义变化检测损失 L_{scd}^1 和 L_{scd}^2 以及变化检测上的对比学习损失函数 L_c 构成;语义变化检测损失 L_{scd}^1 和 L_{scd}^2 为交叉熵损失,描述为:

$$[0016] \quad L_{scd}^1 = -\frac{1}{T} \sum_{t=1}^T y_t^1 \log(p_t^1)$$

$$[0017] \quad L_{scd}^2 = -\frac{1}{T} \sum_{t=1}^T y_t^2 \log(p_t^2)$$

[0018] 其中T为像素的个数, y_t^1 和 p_t^1 , y_t^2 和 p_t^2 分别表示真值标签和解码器 ϕ_{scd}^1 和 ϕ_{scd}^2 在第t个像素上的预测概率;

[0019] 总体的损失函数L描述为:

$$[0020] \quad L = L_{scd}^1 + L_{scd}^2 + L_c$$

[0021] 得到总体的损失函数后进行反向传播,使用AdamW优化器进行优化,重复迭代直至迭代次数达到设置初始值时判定完成训练。

[0022] 本发明进一步的技术方案:步骤1具体如下:

[0023] 步骤1-1:对于 T^1 时相,遥感图像 I^1 通过卷积模块 M_0 得到的维度为 $H_0 \times W_0 \times 64$ 的特征图 m_0^1 , $H_0 = H_{input}/4$ 与 $W_0 = W_{input}/4$,其中 H_{input} 与 W_{input} 为输入图像 I^1 的高和宽;

[0024] 步骤1-2:将 m_0^1 输入高分辨网络编码器 ϕ_e^1 ,高分辨网络编码器 ϕ_e^1 包括4层,第i层包括i个卷积模块 $M_{i,j}$, $i \geq 1, 0 \leq j < i$;卷积模块 $M_{i,j}$ 输出的特征图定义为 $m_{i,j}^1$,特征图 $m_{i,j}^1$ 的分辨率记为 $H_{i,j} \times W_{i,j}$,通道数记为 $C_{i,j}$,其中 $H_{i,j} = H_0/2^j$, $W_{i,j} = W_0/2^j$, $C_{i,j} = 40 \times 2^j$;

[0025] 当 $i \geq 2$ 时,将第i-1层卷积模块 $M_{i-1,y}$ 的输出 $m_{i-1,y}^1$ 经过如下处理统一输入分辨率及通道数后采用加法融合输入第i层的卷积模块 $M_{i,j}$, $y \in [0, i-2]$;

[0026] 当 $y < j$ 时,对特征图 $m_{i-1,y}^1$ 进行j-y次步长为2的跨步卷积,每次跨步卷积通过 3×3 卷积使得特征图的通道数翻倍,分辨率减半;

[0027] 当 $y = j$ 时,对特征图 $m_{i-1,y}^1$ 进行 3×3 卷积进行特征的提取且通道数与分辨率不变;

[0028] 当 $y > j$ 时,对特征图 $m_{i-1,y}^1$ 进行 3×3 卷积,并将通道数变为 40×2^j ,且使用双线性插值上采样操作使得分辨率变为 $H_0/2^j \times W_0/2^j$;

[0029] 通过高分辨编码器中所有的卷积模块后,得到的4个不同尺度的特征图 $m_{4,k}^1$, $k = 0, 1, 2, 3$; $m_{4,k}^1$ 的分辨率与通道数分别为 $H_0/2^k \times W_0/2^k$ 和 40×2^k ;将 $m_{4,k}^1$ 通过上采样操作统一分辨率,再沿通道维度进行拼接,得到语义上下文特征对 f_{sem}^1 和 f_{sem}^2 ;

[0030] 步骤1-3,将语义上下文特征对 f_{sem}^1 和 f_{sem}^2 输入变化特征提取模块,对 f_{sem}^1 和 f_{sem}^2

作绝对差 d_{cd} ,再将 d_{cd} 进行1次 1×1 卷积操作,得到尺度为 $H_0 \times W_0 \times 600$ 变化特征 f_{cd} ;

[0031] 步骤1-4,对于 T^1 时相,将 f_{sem}^1 和 f_{cd} 输入特征融合模块,特征融合模块包含1个沿通道维度的拼接操作和1个改变通道数的 1×1 卷积操作;首先沿通道维度进行拼接,再通过1次 1×1 卷积操作将通道数压缩为原来的1/2,得到尺度为 $H_0 \times W_0 \times 600$ 的语义变化特征 f_{scd}^1 ;

[0032] 步骤1-5,将 f_{scd}^1 输入解码器 φ_{scd}^1 ,解码器包含2个卷积层,一个双线性插值上采样操作和一个softmax操作;第一个卷积层包含1个步幅为1,填充为2的 3×3 卷积,该卷积操作将输入特征向量通道数变为原来1/4,分辨率保持不变,即 $H_0 \times W_0 \times 150$,以及1个批归一化操作和1修正线性单元;第二个卷积层包含1个步幅为1的 1×1 卷积操作,将输入特征向量通道数变为语义变化检测类别数,分辨率保持不变,即 $H_0 \times W_0 \times 7$;双线性插值上采样操作将输入特征向量的分辨率恢复为输入图像的分辨率,即 $H_{input} \times W_{input} \times 7$,最后用softmax操作进行归一化,得到时相的语义变化检测结果 O_{scd}^1 ;

[0033] T^2 时相同理得到 O_{scd}^2 。

[0034] 本发明进一步的技术方案:所有的卷积模块都由 3×3 卷积层、批归一化层与修正线性单元组成。

[0035] 本发明进一步的技术方案:步骤2所述的半难半易的采样策略是对每个类别的锚点,都采样一半的难分类样本和一半的易分类样本,对于每个锚点的负样本,都采用一半的难分类样本和一半的易分类样本;用阈值 δ 划分难分类样本和易分类样本;难分类样本在其对应类别上的二值变化检测预测概率小于阈值 δ ,而易分类样本在其对应类别上的二值变化检测预测概率大于阈值 δ ,其中二值变化检测预测概率来自编码器softmax归一化后的结果。

[0036] 一种计算机系统,其特征在于包括:一个或多个处理器,计算机可读存储介质,用于存储一个或多个程序,其中,当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现上述的方法。

[0037] 一种计算机可读存储介质,其特征在于存储有计算机可执行指令,所述指令在被执行时用于实现上述的方法。

[0038] 有益效果

[0039] 本发明提供了一种基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法;构建了基于高分辨网络的简单可扩展的直接语义变化检测模型来进行语义变化检测,并在变化检测上应用对比学习损失作为监督,有效挖掘由于类别不平衡而导致的难分类样本,提高网络对难分类样本的分类性能;本发明提出的模型得到的在变化区域上细节更加精确,形状更加完整,同时对语义变化类别的分类更加准确。

[0040] 本发明设计了构建一个基于高分辨网络的简单可扩展的直接语义变化检测模型SFSCDNet,并用对比学习损失监督变化检测。在对比学习中,采用半难半易的采样策略,使网络在关注难分类样本的同时易于收敛。通过半难半易的采样策略,对比损失可以指导网络为变化检测提供适当的监督,特别是更加关注难以正确分成变化类或未变化类的样本,从而提高模型的整体语义变化性能。

附图说明

[0041] 附图仅用于示出具体实施例的目的,而并不认为是对本发明的限制,在整个附图中,相同的参考符号表示相同的部件。

[0042] 图1是本发明实施例方法的网络结构图。

[0043] 图2是本发明高分辨编码器的网络结构图。

[0044] 图3是本发明实施例网络模型中对比学习特征表示头结构图。

[0045] 图4是本发明实施例方法和其他现有方法的语义变化检测结果对比图。

具体实施方式

[0046] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图和实施例,对本发明进行进一步详细说明。应当理解,此处所描述的具体实施例仅用以解释本发明,并不用于限定本发明。此外,下面描述的本发明各个实施方式中所涉及到的技术特征只要彼此之间未构成冲突就可以相互组合。

[0047] 本发明提供一种基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法用于高分辨遥感图像的语义变化检测,通过构建一个基于高分辨网络简单可扩展的直接语义变化检测模型和设计对比学习以监督二值变化检测,来解决现有语义变化检测方法的结果精度较低的问题。本发明构建的基于高分辨网络简单可扩展的直接语义变化检测模型用两个共享权重的高分辨网络作为编码器提取双时相遥感图像的语义上下文特征,然后将提取到的双时相语义上下文特征输入变化特征提取模块得到一个变化特征,对于每一个时相,将语义上下文特征和变化特征输入特征融合模块得到语义变化特征,最后将双时相语义变化特征输入两个解码器得到双时相语义变化检测结果。为了使用对比学习监督变化检测,在变化特征提取模块后加入对比学习特征表示头,为对比学习提供变化特征表示,然后用半难半易的采样策略在对比学习提供变化特征表示进行采样以计算对比学习损失函数。双时相语义变化检测损失和变化检测上的对比学习损失构成总体的损失函数,用于优化整体模型。

[0048] 包括如下步骤:

[0049] 步骤1:构建一个基于高分辨网络的简单可扩展的直接语义变化检测模型来进行语义变化检测;

[0050] 将 T^1 和 T^2 时相的遥感图像 I^1 和 I^2 输入两个共享权重的高分辨网络编码器 ϕ_e^1 和 ϕ_e^2 得到语义上下文特征对 f_{sem}^1 和 f_{sem}^2 ,然后将 f_{sem}^1 和 f_{sem}^2 输入变化特征提取模块得到变化特征 f_{cd} 。对于 T^1 时相,将 f_{sem}^1 和 f_{cd} 输入特征融合模块,先对 f_{sem}^1 和 f_{cd} 作绝对差得到 d_{cd} ,再将 d_{cd} 进行1次 1×1 卷积操作得到语义变化特征 f_{sca}^1 ,对于 T^2 时相,将 f_{sem}^2 和 f_{cd} 输入特征融合模块得到语义变化特征 f_{sca}^2 。最后将 f_{sca}^1 和 f_{sca}^2 输入两个解码器 ϕ_{sca}^1 和 ϕ_{sca}^2 得到 T^1 和 T^2 时相的语义变化检测结果 O_{sca}^1 和 O_{sca}^2 ;由于两个时相在编码和解码阶段步骤相同,所以下面仅以 T^1 时相为例, T^2 时相同理;

[0051] 步骤1-1:对于 T^1 时相,遥感图像 I^1 通过卷积模块 M_0 得到的维度为 $H_0 \times W_0 \times 64$ 的特征图 m_0^1 ,其中 H_{input} 与 W_{input} 为输入图像 I^1 的高和宽, $H_0 = H_{input}/4$ 与 $W_0 = W_{input}/4$;

[0052] 步骤1-2:将 m_0^1 输入高分辨网络编码器 ϕ_e^1 ,高分辨网络编码器 ϕ_e^1 包括4层,第i层包括i个卷积模块 $M_{i,j}$, $i \geq 1, 0 \leq j < i$;卷积模块 $M_{i,j}$ 输出的特征图定义为 $m_{i,j}^1$,特征图 $m_{i,j}^1$ 的分辨率记为 $H_{i,j} \times W_{i,j}$,通道数记为 $C_{i,j}$,其中 $H_{i,j} = H_0/2^j, W_{i,j} = W_0/2^j, C_{i,j} = 40 \times 2^j$;

[0053] 当 $i \geq 2$ 时,将第i-1层卷积模块 $M_{i-1,y}$ 的输出 $m_{i-1,y}^1$ 经过如下处理统一输入分辨率及通道数后采用加法融合输入第i层的卷积模块 $M_{i,j}, y \in [0, i-2]$;

[0054] 当 $y < j$ 时,对特征图 $m_{i-1,y}^1$ 进行j-y次步长为2的跨步卷积,每次跨步卷积通过 3×3 卷积使得特征图的通道数翻倍,分辨率减半;

[0055] 当 $y = j$ 时,对特征图 $m_{i-1,y}^1$ 进行 3×3 卷积进行特征的提取且通道数与分辨率不变;

[0056] 当 $y > j$ 时,对特征图 $m_{i-1,y}^1$ 进行 3×3 卷积,并将通道数变为 40×2^j ,且使用双线性插值上采样操作使得分辨率变为 $H_0/2^j \times W_0/2^j$;

[0057] 通过高分辨编码器中所有的卷积模块后,得到的4个不同尺度的特征图 $m_{4,k}^1, k = 0, 1, 2, 3$; $m_{4,k}^1$ 的分辨率与通道数分别为 $H_0/2^k \times W_0/2^k$ 和 40×2^k ;将 $m_{4,k}^1$ 通过上采样操作统一分辨率,再沿通道维度进行拼接,得到语义上下文特征对 f_{sem}^1 和 f_{sem}^2 ;

[0058] 步骤1-3,将语义上下文特征对 f_{sem}^1 和 f_{sem}^2 输入变化特征提取模块,对 f_{sem}^1 和 f_{sem}^2 作绝对差 d_{cd} ,再将 d_{cd} 进行1次 1×1 卷积操作,得到尺度为 $H_0 \times W_0 \times 600$ 变化特征 f_{cd} ;

[0059] 步骤1-4,对于 T^1 时相,将 f_{sem}^1 和 f_{cd} 输入特征融合模块,特征融合模块包含1个沿通道维度的拼接操作和1个改变通道数的 1×1 卷积操作。首先沿通道维度进行拼接,再通过1次 1×1 卷积操作将通道数压缩为原来的1/2,得到尺度为 $H_0 \times W_0 \times 600$ 的语义变化特征 f_{scd}^1 ;

[0060] 步骤1-5,将 f_{scd}^1 输入解码器 ϕ_{scd}^1 ,解码器包含2个卷积层,一个双线性插值上采样操作和一个softmax操作。第一个卷积层包含1个步幅为1,填充为2的 3×3 卷积,该卷积操作将输入特征向量通道数变为原来1/4,分辨率保持不变,即 $H_0 \times W_0 \times 150$,以及1个批归一化操作和1修正线性单元;第二个卷积层包含1个步幅为1的 1×1 卷积操作,将输入特征向量通道数变为语义变化检测类别数,分辨率保持不变,即 $H_0 \times W_0 \times 7$;双线性插值上采样操作将输入特征向量的分辨率恢复为输入图像的分辨率,即 $H_{input} \times W_{input} \times 7$,最后用softmax操作进行归一化,得到时相的语义变化检测结果 O_{scd}^1 ;

[0061] T^2 时相同理;

[0062] 优选地,所有的卷积模块都由 3×3 卷积层、批归一化层与修正线性单元组成。

[0063] 步骤2,使用对比学习损失监督变化检测。

[0064] 在变化特征提取模块后加入对比学习特征表示头 Φ_{rep} ,将变化特征 f_{cd} 输入对比学习特征表示头 Φ_{rep} 得到密集变化特征表示 Z_{rep} , Φ_{rep} 包含2个卷积层。第一个卷积层包含1个步幅为1,填充为2的 3×3 卷积,该卷积操作将输入变化特征 f_{cd} 通道数变为原来1/4,分辨率保持不变,即 $H_0 \times W_0 \times 150$,以及1个批归一化操作和1修正线性单元;第二个卷积层包含1个步幅为1的 1×1 卷积操作,将输入特征向量通道数变为256,分辨率保持不变,即 $H_0 \times W_0 \times$

256; 由于对比损失仅在训练期间提供监督, 因此对比学习特征表示头在推理期间被移除; 用半难半易的采样策略在变化特征表示 Z 中对变化类别和未变化类别的样本对应的特征向量进行采样以计算对比学习损失函数 L_c :

$$[0065] \quad L_c = -\frac{1}{C \times A} \sum_{c=0}^{C-1} \sum_{i=1}^A \log \left[\frac{e^{\langle z_{ca}, z_{ca}^+ \rangle / \tau}}{e^{\langle z_{ca}, z_{ca}^+ \rangle / \tau} + \sum_{b=1}^B e^{\langle z_{ca}, z_{ca,b}^- \rangle / \tau}} \right]$$

[0066] 其中, 由于类别为变化类别和未变化类别, 所以 $\|C\|=2$ 。 z_{ca} 是第 c 类第 a 个锚点的特征向量, z_{ca}^+ 是第 c 类的正样本, 是所有第 c 类样本特征向量的平均值, $z_{ca,b}^-$ 是第 c 类第 a 个锚点的第 b 个属于其他类别的负样本的特征向量。对于 C 中的每个类别, 锚点和锚点的负样本对应的特征向量都采集自变化特征表示 Z_{rep} , 每个类别都有 A 个锚点, 并且每个锚点都有一个正样本和 B 个负样本。此处, $A=512$, $B=512$ 。 $\langle \cdot, \cdot \rangle$ 是两个特征向量之间的余弦相似度, 用于度量两个特征向量之间的距离, 范围为 -1 到 1 , $\tau=0.5$ 是温度系数。对于当前训练批量中的每个类别, 通过优化网络参数使得 L_c 最小, 来拉近该类别的锚点与其正样本的距离, 而推远该类别的锚点与其负样本的距离。

[0067] 半难半易的采样策略是对每个类别的锚点, 都采样一半的难分类样本和一半的易分类样本, 对于每个锚点的负样本, 都采用一半的难分类样本和一半的易分类样本。用阈值 δ 划分难分类样本和易分类样本。难分类样本在其对应类别上的二值变化检测预测概率小于阈值 δ , 而易分类样本在其对应类别上的二值变化检测预测概率大于阈值 δ , 其中二值变化检测预测概率来自编码器softmax归一化后的结果。难分类样本引导网络更关注难以分类的样本, 但是过多的难分类样本会使网络难以收敛。因此, 通过选择一半难分类样本和一半易分类样本, 可以使网络在关注难分类样本的同时易于收敛。通过半难半易的采样策略, 对比损失可以指导网络为变化检测提供适当的监督, 特别是更加关注难以正确分成变化类或未变化类的样本, 从而提高模型的整体语义变化性能。

[0068] 步骤3, 通过使总体的损失函数 L 最小来优化模型的参数, 总体的损失函数 L 由 T^1 和 T^2 的语义变化检测损失 L_{scd}^1 和 L_{scd}^2 以及变化检测上的对比学习损失函数 L_c 构成。语义变化检测损失 L_{scd}^1 和 L_{scd}^2 为交叉熵损失, 描述为:

$$[0069] \quad L_{scd}^1 = -\frac{1}{T} \sum_{t=1}^T y_t^1 \log(p_t^1)$$

$$[0070] \quad L_{scd}^2 = -\frac{1}{T} \sum_{t=1}^T y_t^2 \log(p_t^2)$$

[0071] 其中 T 为像素的个数, y_t^1 和 p_t^1 , y_t^2 和 p_t^2 分别表示真值标签和解码器 φ_{scd}^1 和 φ_{scd}^2 在第 t 个像素上的预测概率。

[0072] 总体的损失函数 L 描述为:

$$[0073] \quad L = L_{scd}^1 + L_{scd}^2 + L_c$$

[0074] 得到总体的损失函数后进行反向传播, 使用AdamW优化器进行优化, 重复迭代直至迭代次数达到设置初始值时判定完成训练。

[0075] 实施例:

[0076] 如图1所示,为基于二值变化检测对比学习的高分辨遥感图像语义变化检测方法的模型。其包括高分辨率遥感图像语义变化检测模型SFSCDNet和在变化检测上的对比学习算法。其中高分辨编码器的结构如图2所示,对比学习特征表示头的结构如图3所示。具体方法包括以下步骤:

[0077] S1,将 T^1 和 T^2 时相的遥感图像 I^1 和 I^2 输入两个共享权重的高分辨网络编码器 ϕ_e^1 和 ϕ_e^2 得到语义上下文特征对 f_{sem}^1 和 f_{sem}^2 ,然后将 f_{sem}^1 和 f_{sem}^2 输入变化特征提取模块得到变化特征 f_{cd} 。对于 T^1 时相,将 f_{sem}^1 和 f_{cd} 输入特征融合模块,先对 f_{sem}^1 和 f_{sem}^2 作绝对差得到 d_{cd} ,再将 d_{cd} 进行1次 1×1 卷积操作得到语义变化特征 f_{scd}^1 ,对于 T^2 时相,将 f_{sem}^2 和 f_{cd} 输入特征融合模块得到语义变化特征 f_{scd}^2 。最后将 f_{scd}^1 和 f_{scd}^2 输入两个解码器 ϕ_{scd}^1 和 ϕ_{scd}^2 得到 T^1 和 T^2 时相的语义变化检测结果 O_{scd}^1 和 O_{scd}^2 ;

[0078] S2,在变化特征提取模块后加入对比学习特征表示头 Φ_{rep} ,将变化特征 f_{cd} 输入对比学习特征表示头 Φ_{rep} 得到密集变化特征表示 Z_{rep} ,使用对比学习损失监督变化检测;

[0079] S3,用AdamW优化器最小化总体的损失函数L来优化模型参数,L由 T^1 和 T^2 的语义变化检测损失 L_{scd}^1 和 L_{scd}^2 以及变化检测上的对比学习损失函数 L_c 构成。

[0080] 本实施例中,步骤S1的执行网络简称为SFSCDNet。下文将结合SFSCDNet的结构,对步骤S1-步骤S3的执行过程进一步详细说明。

[0081] 本实施例中,参见图1和图2,步骤S1将 T^1 和 T^2 时相分辨率为 $H_{input} \times W_{input} \times 3$ 的遥感图像 I^1 和 I^2 输入共享权重的在ImageNet上预训练的高分辨网络编码器 ϕ_e^1 和 ϕ_e^2 ,得到分辨率为 $H_{input}/4 \times W_{input}/4 \times 600$ 的语义上下文特征对 f_{sem}^1 和 f_{sem}^2 。变化特征提取模块对 f_{sem}^1 和 f_{sem}^2 作绝对差 d_{cd} ,再将 d_{cd} 进行1次 1×1 卷积操作,得到分辨率为 $H_{input}/4 \times W_{input}/4 \times 600$ 变化特征 f_{cd} 。对于 T^1 时相,将 f_{sem}^1 和 f_{cd} 输入特征融合模块,首先沿通道维度进行拼接,再通过1次 1×1 卷积操作将通道数压缩为原来的1/2,得到分辨率为 $H_{input}/4 \times W_{input}/4 \times 600$ 的语义变化特征 f_{scd}^1 , T^2 时相同理,经过特征融合模块,得到分辨率为 $H_{input}/4 \times W_{input}/4 \times 600$ 的语义变化特征 f_{scd}^2 。将 f_{scd}^1 和 f_{scd}^2 输入两个结构相同但不共享权重的解码器 ϕ_{scd}^1 和 ϕ_{scd}^2 ,解码器包含2个卷积层,一个双线性插值上采样操作和一个softmax操作。第一个卷积层包含1个步幅为1,填充为2的 3×3 卷积,该卷积操作将输入特征向量通道数变为原来1/4,分辨率保持不变,即 $H_{input}/4 \times W_{input}/4 \times 150$,以及包含1个批归一化操作和1修正线性单元;第二个卷积层包含1个步幅为1的 1×1 卷积操作,将输入特征向量通道数变为语义变化检测类别数,即 $H_{input}/4 \times W_{input}/4 \times 7$,分辨率保持不变;双线性插值上采样操作将输入特征向量的分辨率恢复为输入图像的分辨率,最后用softmax操作进行归一化,得到 T^1 和 T^2 时相的语义变化检测结果 O_{scd}^1 和 O_{scd}^2 。

[0082] 为了监督语义变化检测中的变化检测,步骤S2在变化特征提取模块后加入对比学习特征表示头 Φ_{rep} ,用对比学习损失函数监督变化检测,将变化特征 f_{cd} 输入对比学习特征

表示头 Φ_{rep} 得到密集语义特征表示 Z_{rep} , Φ_{rep} 的结构如图3所示。 Φ_{rep} 包含2个卷积层, 第一个卷积层包含1个步幅为1, 填充为2的 3×3 卷积, 该卷积操作将变化特征 f_{cd} 通道数变为原来 $1/4$, 分辨率保持不变, 即 $H_0 \times W_0 \times 150$, 以及1个批归一化操作和1修正线性单元; 第二个卷积层包含1个步幅为1的 1×1 卷积操作, 将输入特征向量通道数变为256, 分辨率保持不变, 即 $H_0 \times W_0 \times 256$ 。由于对比损失仅在训练期间提供监督, 因此对比学习特征表示头在推理期间被移除。

[0083] 步骤S2用半难半易的采样策略在变化特征表示 Z_{rep} 中对变化类别和未变化类别的样本对应的特征向量进行采样以计算对比学习损失函数 L_c :

$$[0084] \quad L_c = -\frac{1}{C \times A} \sum_{c=0}^{C-1} \sum_{i=1}^A \log \left[\frac{e^{\langle z_{ca}, z_{ca}^+ \rangle / \tau}}{e^{\langle z_{ca}, z_{ca}^+ \rangle / \tau} + \sum_{b=1}^B e^{\langle z_{ca}, z_{ca,b}^- \rangle / \tau}} \right]$$

[0085] 其中, 由于类别为变化类别和未变化类别, 所以 $\|C\|=2$ 。 z_{ca} 是第 c 类第 a 个锚点的特征向量, z_{ca}^+ 是第 c 类的正样本, 是所有第 c 类样本特征向量的平均值, $z_{ca,b}^-$ 是第 c 类第 a 个锚点的第 b 个属于其他类别的负样本的特征向量。对于 C 中的每个类别, 锚点和锚点的负样本对应的特征向量都采集自变化特征表示 Z_{rep} , 每个类别都有 A 个锚点, 并且每个锚点都有一个正样本和 B 个负样本。此处, $A=512$, $B=512$ 。 $\langle \cdot, \cdot \rangle$ 是两个特征向量之间的余弦相似度, 用于度量两个特征向量之间的距离, 范围为 -1 到 1 , $\tau=0.5$ 是温度系数。对于当前训练批量中的每个类别, 通过优化网络参数使得 L_c 最小, 来拉近该类别的锚点与其正样本的距离, 而推远该类别的锚点与其负样本的距离。

[0086] 半难半易的采样策略是对每个类别的锚点, 都采样一半的难分类样本和一半的易分类样本, 对于每个锚点的负样本, 都采用一半的难分类样本个一半的易分类样本。用阈值 δ 划分难分类样本和易分类样本。难分类样本在其对应类别上的二值变化检测预测概率小于阈值 δ , 而易分类样本在其对应类别上的二值变化检测预测概率大于阈值 δ , 其中二值变化检测预测概率来自编码器 softmax 归一化后的结果。难分类样本引导网络更关注难以分类的样本, 但是过多的难分类样本会使网络难以收敛。因此, 通过选择一半难分类样本和一半易分类样本, 可以使网络在关注难分类样本的同时易于收敛。通过半难半易的采样策略, 对比损失可以指导网络为变化检测提供适当的监督, 特别是更加关注难以正确分成变化类或未变化类的样本, 从而提高模型的整体语义变化性能。

[0087] 步骤S3总体的损失函数 L 描述为:

$$[0088] \quad L = L_{sca}^1 + L_{sca}^2 + L_c$$

[0089] 其中语义变化检测损失 L_{sca}^1 和 L_{sca}^2 为交叉熵损失, 描述为:

$$[0090] \quad L_{sca}^1 = -\frac{1}{T} \sum_{t=1}^T y_t^1 \log(p_t^1)$$

$$[0091] \quad L_{sca}^2 = -\frac{1}{T} \sum_{t=1}^T y_t^2 \log(p_t^2)$$

[0092] 其中 T 为像素的个数, y_t^1 和 p_t^1 , y_t^2 和 p_t^2 分别表示真值标签和解码器 φ_{sca}^1 和 φ_{sca}^2 在

第t个像素上的预测概率。

[0093] 得到总体的损失函数后进行反向传播,使用AdamW优化器进行优化,重复迭代直至迭代次数达到设置初始值时判定完成训练。

[0094] 为了验证本方法的有效性,本实施例使用公开数据集SECOND进行了网络框架的训练和测试,并且与其他方法进行了对比。SECOND数据集包含2968组数据,每一组数据包含两张不同时相的图像,每张图像大小为 512×512 ,并且这2968组数据集都包含了变化区域。按照9:1的比例划分训练集和测试集。

[0095] 本实例提出的算法与7种最新的变化检测方法,DSCD、SCDS、ICDS、ChangeMask、HBSCD、Bi-SRNet和SCDNet进行对比,具体结果如表1所示。评价指标共有3种,分别为mIoU、SeK和Score。结合表1可以看出,本实例实施方法3个评价指标,在mIoU、SeK和Score上达到73.83%、26.37%和40.61%,均是最优结果。与第二好的SCDNet相比,本方法将mIoU提高了0.77%,SeK提高了2.71%,Score提高了2.02%。图4为本实施例方法和其他现有方法的三组语义变化检测结果对比示意图,通过图4第一行、第二行和第三行可以看出,本实施例方法的在水面和树木上的语义变化检测效果非常接近真实情况,语义变化区域预测完整,且轮廓清晰,但是其他对比方法有误报或者漏检的情况。

[0096] 表1本发明实施例方法和其他现有方法的测试结果对比表

[0097]

Methods	mIoU(%)	Sek(%)	Score(%)
DSCD	62.45	10.20	25.88
SCDS	69.18	14.96	31.22
ICDS	71.95	21.83	36.86
ChangeMask	-	17.89	-
HBSDC	72.40	21.46	36.74
Bi-SRNet	73.41	23.22	38.59
SCDNet	73.06	23.66	38.59
SFSCDNet+对比学习	73.83	26.37	40.61

[0098] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明公开的技术范围内,可轻易想到各种等效的修改或替换,这些修改或替换都应涵盖在本发明的保护范围之内。

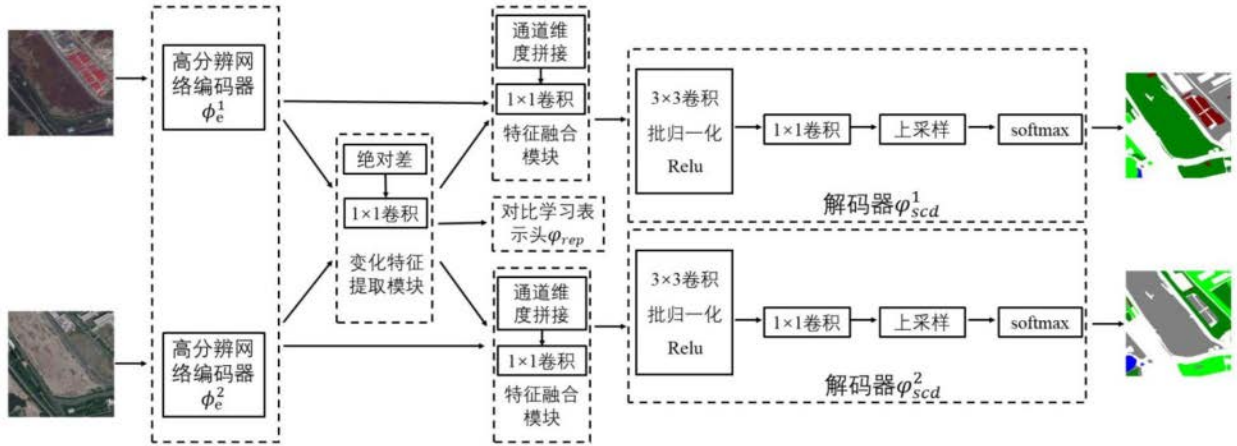


图1

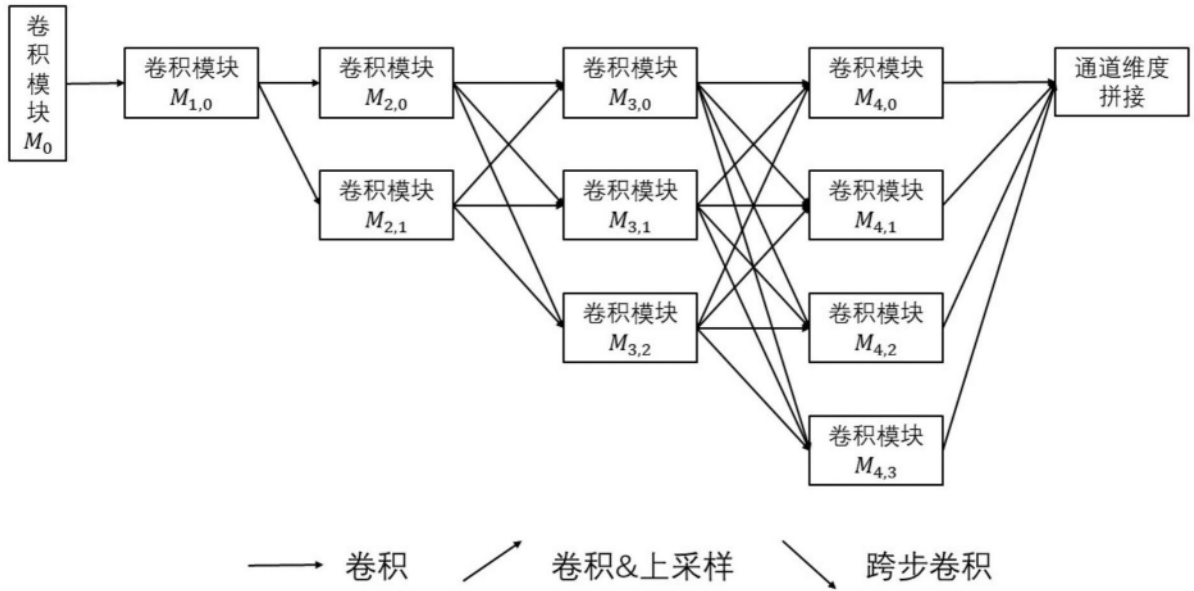


图2

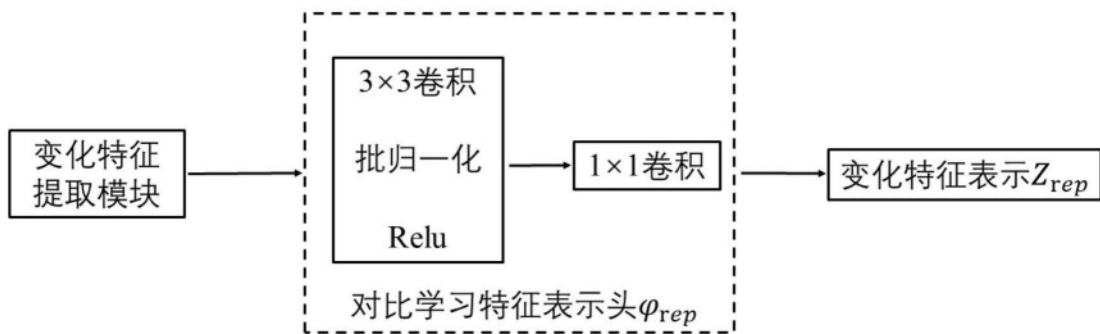


图3

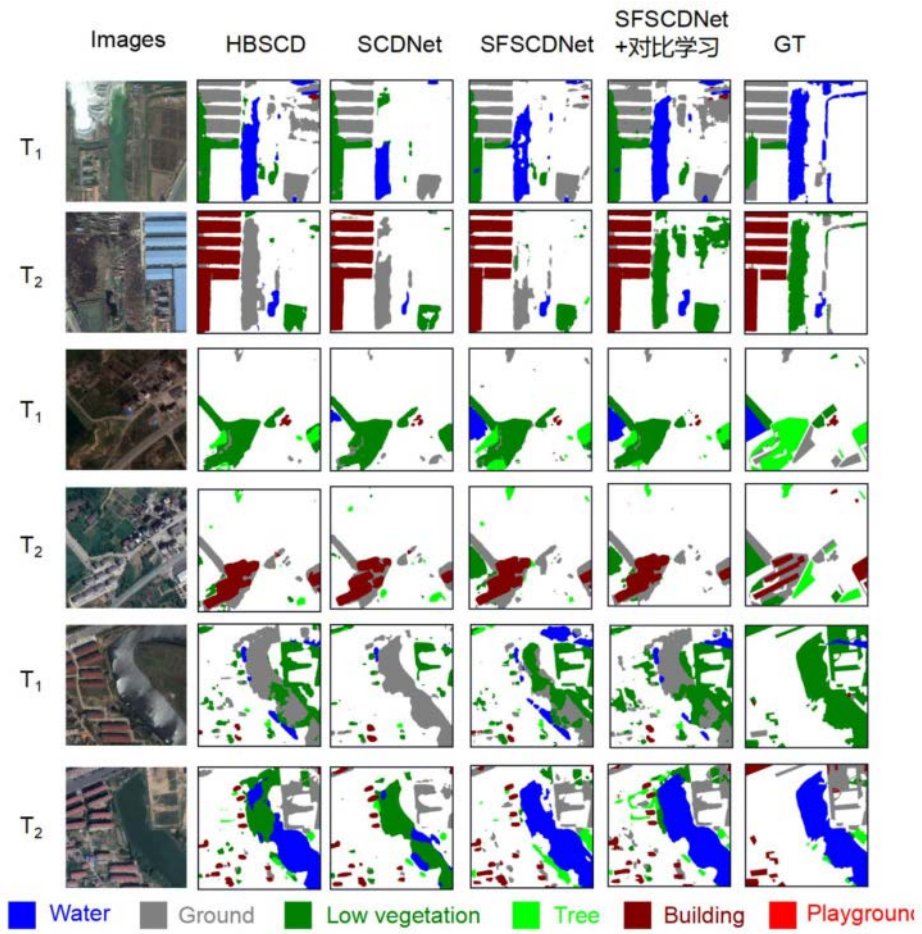


图4