



(12) 发明专利

(10) 授权公告号 CN 113673530 B

(45) 授权公告日 2023.06.16

(21) 申请号 202110948818.6

(22) 申请日 2021.08.18

(65) 同一申请的已公布的文献号
申请公布号 CN 113673530 A

(43) 申请公布日 2021.11.19

(73) 专利权人 中国人民解放军国防科技大学
地址 410073 湖南省长沙市开福区德雅路
109号

(72) 发明人 陈浩 资文杰 熊伟 彭双 杜春
伍江江 吴焯 李军 欧阳雪
杨岸然

(74) 专利代理机构 长沙国科天河知识产权代理
有限公司 43225
专利代理师 邱轶

(51) Int.Cl.

G06V 20/70 (2022.01)

G06V 20/10 (2022.01)

G06V 10/26 (2022.01)

G06V 10/82 (2022.01)

G06V 10/80 (2022.01)

G06N 3/045 (2023.01)

G06N 3/0464 (2023.01)

G06N 3/08 (2023.01)

(56) 对比文件

CN 112132193 A, 2020.12.25

CN 112991351 A, 2021.06.18

CN 11311835 A, 2021.07.13

US 2021174513 A1, 2021.06.10

审查员 马雅凡

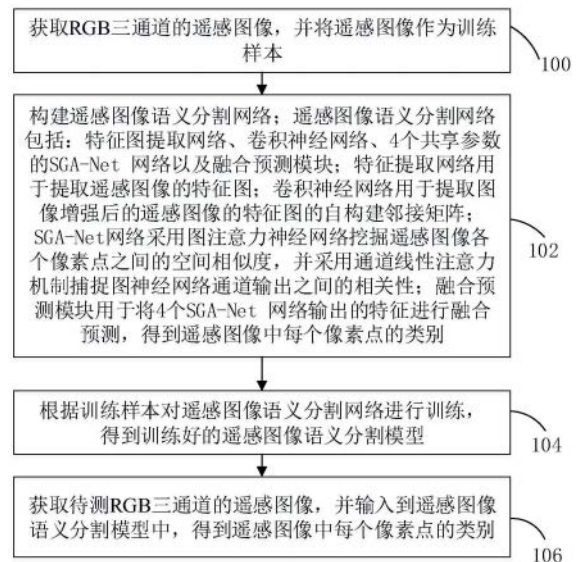
权利要求书2页 说明书10页 附图3页

(54) 发明名称

遥感图像语义分割方法、装置、计算机设备和存储介质

(57) 摘要

本申请涉及一种遥感图像语义分割方法、装置、计算机设备和存储介质。所述方法包括：获取RGB三通道的遥感图像；构建遥感图像语义分割网络，该网络包括：特征图提取网络、卷积神经网络、4个共享参数的SGA-Net网络以及融合预测模块；利用RGB三通道的遥感图像对该网络进行训练，得到遥感图像语义分割模型，利用该模型对待测RGB三通道的遥感图像进行分割，得到遥感图像中每个像素点的类别。采用图神经网络充分地挖掘了遥感图像各个像素点之间的空间相似度，可体现全局的空间联系，采用通道线性注意力机制来捕捉图神经网络通道输出之间的相关性，并进一步提高基于GNN模型的性能。采用本方法可以提高建筑物、道路、低植被的分割精度。



1. 一种遥感图像语义分割方法,其特征在于,所述方法包括:

获取RGB三通道的遥感图像,并将所述遥感图像作为训练样本;

构建遥感图像语义分割网络;所述遥感图像语义分割网络包括:特征图提取网络、卷积神经网络、4个共享参数的SGA-Net网络以及融合预测模块;所述特征图提取网络用于提取遥感图像的特征图;所述卷积神经网络用于从将遥感图像的特征图进行图像增强得到的4张增强特征图中提取出4个自构建邻接矩阵;所述SGA-Net网络包括:图注意力神经网络和通道线性注意力机制;所述SGA-Net网络用于采用图注意力神经网络挖掘遥感图像各个像素点之间的空间相似度,并采用通道线性注意力机制捕捉图神经网络通道输出之间的相关性;4个共享参数的SGA-Net网络,用于将第一个所述自构建邻接矩阵输入到第一个所述SGA-Net网络的图注意力神经网络中,得到全局空间相似特征图;将所述全局空间相似特征图输入到第一个所述SGA-Net网络的通道线性注意力机制中,得到第一个通道相关性特征图;将第二个所述自构建邻接矩阵输入到第二个所述SGA-Net网络中,得到第二个通道相关性特征图;将第三个所述自构建邻接矩阵输入到第三个所述SGA-Net网络中,得到第三个通道相关性特征图;将第4个所述自构建邻接矩阵输入到第4个所述SGA-Net网络中,得到第4个通道相关性特征图;所述融合预测模块用于将4个通道相关性特征图进行融合预测,得到遥感图像中每个像素点的预测类别;

根据所述训练样本对所述遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型;

获取待测RGB三通道的遥感图像,并输入到所述遥感图像语义分割模型中,得到遥感图像中每个像素点的类别。

2. 根据权利要求1所述的方法,其特征在于,根据所述训练样本对所述遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型,步骤前还包括:

构建总损失函数,所述总损失函数包括:对角对数正则化项、Kullback-Leibler散度以及自适应多类加权损失函数。

3. 根据权利要求2所述的方法,其特征在于,根据所述训练样本对所述遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型,包括:

将所述训练样本输入到所述遥感图像语义分割网络中得到遥感图像中每个像素点的预测类别;

根据所述预测类别、所述训练样本以及所述总损失函数,对所述遥感图像语义分割网络进行反向训练,得到训练好的遥感图像语义分割模型。

4. 根据权利要求1所述的方法,其特征在于,将所述全局空间相似特征图输入到第一个所述SGA-Net网络的通道线性注意力机制中,得到第一个通道相关性特征图,包括:

将所述全局空间相似特征图输入到所述通道线性注意力机制中,分别对所述全局空间相似特征图进行3次变形操作,得到第一变形特征、第二变形特征以及第三变形特征;

对所述全局空间相似特征图进行变形操作,并对变形得到的结果进行转置,得到转置变形特征;

将所述第三变形特征与所述转置变形特征相乘,并将得到的特征进行变形操作,得到特征矩阵;

将所述特征矩阵与所述第二变形特征进行相乘,并将得到的特征进行变形操作后与所

述第一变形特征相加,得到第一个通道相关性特征图。

5. 根据权利要求1所述的方法,其特征在于,所述图注意力神经网络是8头图注意力神经网络。

6. 根据权利要求1所述的方法,其特征在于,将所述遥感图像的特征图旋转 0° 、 90° 、 180° 以及 270° ,得到4张增强特征图。

7. 一种遥感图像语义分割装置,其特征在于,所述装置包括:

遥感图像获取模块,用于获取RGB三通道的遥感图像,并将所述遥感图像作为训练样本;

遥感图像语义分割网络构建模块,用于构建遥感图像语义分割网络;所述遥感图像语义分割网络包括:特征图提取网络、卷积神经网络、4个共享参数的SGA-Net网络以及融合预测模块;所述特征图提取网络用于提取遥感图像的特征图;所述卷积神经网络用于从将遥感图像的特征图进行图像增强得到的4张增强特征图中提取出4个自构建邻接矩阵;所述SGA-Net网络包括:图注意力神经网络和通道线性注意力机制;所述SGA-Net网络用于采用图注意力神经网络挖掘遥感图像各个像素点之间的空间相似度,并采用通道线性注意力机制捕捉图神经网络通道输出之间的相关性;4个共享参数的SGA-Net网络,用于将第一个所述自构建邻接矩阵输入到第一个所述SGA-Net网络的图注意力神经网络中,得到全局空间相似特征图;将所述全局空间相似特征图输入到第一个所述SGA-Net网络的通道线性注意力机制中,得到第一个通道相关性特征图;将第二个所述自构建邻接矩阵输入到第二个所述SGA-Net网络中,得到第二个通道相关性特征图;将第三个所述自构建邻接矩阵输入到第三个所述SGA-Net网络中,得到第三个通道相关性特征图;将第4个所述自构建邻接矩阵输入到第4个所述SGA-Net网络中,得到第4个通道相关性特征图;所述融合预测模块用于将4个通道相关性特征图进行融合预测,得到遥感图像中每个像素点的预测类别;

遥感图像语义分割网络训练模块,用于根据所述训练样本对所述遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型;

遥感图像像素点类别确定模块,用于获取待测RGB三通道的遥感图像,并输入到所述遥感图像语义分割模型中,得到遥感图像中每个像素点的类别。

8. 一种计算机设备,包括存储器和处理器,所述存储器存储有计算机程序,其特征在于,所述处理器执行所述计算机程序时实现权利要求1至6中任一项所述方法的步骤。

9. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1至6中任一项所述的方法的步骤。

遥感图像语义分割方法、装置、计算机设备和存储介质

技术领域

[0001] 本申请涉及遥感图像处理技术领域,特别是涉及一种遥感图像语义分割方法、装置、计算机设备和存储介质。

背景技术

[0002] 遥感图像语义分割一直是计算机视觉的研究重点,遥感图像的语义分割是地面物体解读中的一个迫切问题,其目旨在为图像中的每个像素分配一个明确的物体类别。它已经成为交通监测、交通安全、交通监测、环境保护、车辆检测等方面最重要的方法之一。遥感图像通常由各种物体、高度不平衡的地面和错综复杂的色彩变化的纹理组成,给遥感图像的语义分割带来了挑战和困难。在深度学习时代之前,为了显示植被和土地覆盖的分布,超级像素经常被用来作为从多光谱图像中提取特征的措施。然而,手工制作的描述符对于模拟这些指数的灵活性来说是一个挑战。

[0003] 现有的基于卷积神经网络的遥感图像语义分割的方法模型只可以从有限的感受野而不是整个图像中捕获像素的相互依赖性;基于图神经网络的遥感影像语义分割的方法模型,需要通过先验知识得到的邻接矩阵,泛化性不强。

发明内容

[0004] 基于此,有必要针对上述技术问题,提供一种能够挖掘遥感图像的全局像素点之间空间相似性的遥感图像语义分割方法、装置、计算机设备和存储介质。

[0005] 一种遥感图像语义分割方法,所述方法包括:

[0006] 获取RGB三通道的遥感图像,并将所述遥感图像作为训练样本。

[0007] 构建遥感图像语义分割网络;所述遥感图像语义分割网络包括:特征图提取网络、卷积神经网络、4个共享参数的SGA-Net网络以及融合预测模块;所述特征提取网络用于提取遥感图像的特征图;所述卷积神经网络用于提取图像增强后的遥感图像的特征图的自构建邻接矩阵;所述SGA-Net网络采用图注意力神经网络挖掘遥感图像各个像素点之间的空间相似度,并采用通道线性注意力机制捕捉图神经网络通道输出之间的相关性;所述融合预测模块用于将4个SGA-Net网络输出的特征进行融合预测,得到遥感图像中每个像素点的类别。

[0008] 根据所述训练样本对所述遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型。

[0009] 获取待测RGB三通道的遥感图像,并输入到所述遥感图像语义分割模型中,得到遥感图像中每个像素点的类别。

[0010] 在其中一个实施例中,根据所述训练样本对所述遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型,步骤前还包括:

[0011] 构建总损失函数,所述总损失函数包括:对角对数正则化项、Kullback-Leibler散度以及自适应多类加权损失函数。

[0012] 在其中一个实施例中,根据所述训练样本对所述遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型,包括:

[0013] 将所述训练样本输入到所述特征图提取网络中,得到遥感图像的特征图。

[0014] 将所述遥感图像的特征图进行图像增强,得到4张增强特征图。

[0015] 将4张所述增强特征图输入到所述卷积神经网络中,得到4个自构建邻接矩阵。

[0016] 将4个所述自构建邻接矩阵分别输入到4个所述共享参数的SGA-Net网络得到4个通道相关性特征图。

[0017] 将4个所述通道注意力特征图输入到所述融合预测网络中,得到遥感图像中每个像素点的预测类别。

[0018] 根据所述预测类别、所述训练样本以及所述总损失函数,对所述遥感图像语义分割网络进行反向训练,得到训练好的遥感图像语义分割模型。

[0019] 在其中一个实施例中,所述SGA-Net网络包括:图注意力神经网络和通道线性注意力机制。

[0020] 将4个所述自构建邻接矩阵分别输入到4个所述共享参数的SGA-Net网络得到4个通道相关性特征图,包括:

[0021] 将第一个所述自构建邻接矩阵输入到第一个所述SGA-Net网络的图注意力神经网络中,得到全局空间相似特征图。

[0022] 将所述全局空间相似特征图输入到第一个所述SGA-Net网络的通道线性注意力机制中,得到第一个通道相关性特征图。

[0023] 将第二个所述自构建邻接矩阵输入到第二个所述SGA-Net网络中,得到第二个通道相关性特征图。

[0024] 将第三个所述自构建邻接矩阵输入到第三个所述SGA-Net网络中,得到第三个通道相关性特征图。

[0025] 将第4个所述自构建邻接矩阵输入到第4个所述SGA-Net网络中,得到第4个通道相关性特征图。

[0026] 在其中一个实施例中,将所述全局空间相似特征图输入到第一个所述SGA-Net网络的通道线性注意力机制中,得到第一个通道相关性特征图,包括:

[0027] 将所述全局空间相似特征图输入到所述通道线性注意力机制中,分别对所述全局空间相似特征图进行3次变形操作,得到第一变形特征、第二变形特征以及第三变形特征。

[0028] 对所述全局空间相似特征图进行变形操作,并对变形得到的结果进行转置,得到转置变形特征。

[0029] 将所述第三变形特征与所述转置变形特征相乘,并将得到的特征进行变形操作,得到特征矩阵。

[0030] 将所述特征矩阵与所述第二变形特征进行相乘,并将得到的特征进行变形操作后与所述第一变形特征相加,得到第一个通道相关性特征图。

[0031] 在其中一个实施例中,所述图注意力神经网络是8头图注意力神经网络。

[0032] 在其中一个实施例中,将所述遥感图像的特征图进行图像增强,得到4张增强特征图,包括:

[0033] 将所述遥感图像的特征图旋转 0° 、 90° 、 180° 以及 270° ,得到4张增强特征图。

[0034] 一种遥感图像语义分割装置,所述装置包括:

[0035] 遥感图像获取模块,用于获取RGB三通道的遥感图像,并将所述遥感图像作为训练样本。

[0036] 遥感图像语义分割网络构建模块,用于构建遥感图像语义分割网络;所述遥感图像语义分割网络包括:特征图提取网络、卷积神经网络、4个共享参数的SGA-Net网络以及融合预测模块;所述特征提取网络用于提取遥感图像的特征图;所述卷积神经网络用于提取图像增强后的遥感图像的特征图的自构建邻接矩阵;所述SGA-Net网络采用图注意力神经网络挖掘遥感图像各个像素点之间的空间相似度,并采用通道线性注意力机制捕捉图神经网络通道输出之间的相关性;所述融合预测模块用于将4个SGA-Net网络输出的特征进行融合预测,得到遥感图像中每个像素点的类别。

[0037] 遥感图像语义分割网络训练模块,用于根据所述训练样本对所述遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型。

[0038] 遥感图像像素点类别确定模块,用于获取待测RGB三通道的遥感图像,并输入到所述遥感图像语义分割模型中,得到遥感图像中每个像素点的类别。

[0039] 上述遥感图像语义分割方法、装置、计算机设备和存储介质,获取RGB三通道的遥感图像,构建遥感图像语义分割网络;遥感图像语义分割网络包括:特征图提取网络、卷积神经网络、4个共享参数的SGA-Net网络以及融合预测模块;利用RGB三通道的遥感图像对遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型,利用遥感图像语义分割模型对待测RGB三通道的遥感图像进行分割,得到遥感图像中每个像素点的类别。采用图神经网络充分地挖掘了遥感图像各个像素点之间的空间相似度,可以体现全局的空间联系,采用通道线性注意机制来捕捉图神经网络通道输出之间的相关性,并进一步提高基于GNN的模型的性能。采用本方法可以提高建筑物、道路、低植被的分割精度。

附图说明

[0040] 图1为一个实施例中遥感图像语义分割方法的流程示意图;

[0041] 图2为一个实施例中遥感图像语义分割网络的结构示意图;

[0042] 图3为一个实施例中遥感图像语义分割装置的结构框图;

[0043] 图4为一个实施例中计算机设备的内部结构图。

具体实施方式

[0044] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处描述的具体实施例仅仅用以解释本申请,并不用于限定本申请。

[0045] 自构建图注意力神经网络:Self-constructing graph attention neural network,简称:SGA-Net网络。

[0046] 在一个实施例中,如图1所示,提供了一种遥感图像语义分割方法,该方法包括以下步骤:

[0047] 步骤100:获取RGB三通道的遥感图像,并将遥感图像作为训练样本。

[0048] RGB三通道的遥感图像是任意大小的,可以是建筑物、道路或低植被等场景的遥感

图像。

[0049] 步骤102:构建遥感图像语义分割网络。

[0050] 遥感图像语义分割网络包括:特征图提取网络、卷积神经网络、4个共享参数的SGA-Net网络以及融合预测模块。

[0051] 特征提取网络用于提取遥感图像的特征图。

[0052] 卷积神经网络用于提取图像增强后的遥感图像的特征图的自构建邻接矩阵。

[0053] SGA-Net网络采用图注意力神经网络挖掘遥感图像各个像素点之间的空间相似度,并采用通道线性注意力机制捕捉图神经网络通道输出之间的相关性;融合预测模块用于将4个SGA-Net网络输出的特征进行融合预测,得到遥感图像中每个像素点的类别。

[0054] 步骤104:根据训练样本对遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型。

[0055] 步骤106:获取待测RGB三通道的遥感图像,并输入到遥感图像语义分割模型中,得到遥感图像中每个像素点的类别。

[0056] 遥感图像语义分割的目标是给每一个像素分配一个类别。

[0057] 上述遥感图像语义分割方法中,获取RGB三通道的遥感图像,构建遥感图像语义分割网络;遥感图像语义分割网络包括:特征图提取网络、卷积神经网络、4个共享参数的SGA-Net网络以及融合预测模块;利用RGB三通道的遥感图像对遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型,利用遥感图像语义分割模型对待测RGB三通道的遥感图像进行分割,得到遥感图像中每个像素点的类别。采用图神经网络充分地挖掘了遥感图像各个像素点之间的空间相似度,可以体现全局的空间联系,采用通道线性注意机制来捕捉图神经网络通道输出之间的相关性,并进一步提高基于GNN的模型的性能。采用本方法可以提高建筑物、道路、低植被的分割精度。

[0058] 在其中一个实施例中,步骤104前还包括:构建总损失函数,总损失函数包括:对角对数正则化项、Kullback-Leibler散度以及自适应多类加权损失函数。

[0059] 图2所示的遥感图像语义分割网络框架的损失函数是三部分组成的:对角对数正则化项、Kullback-Leibler散度以及自适应多类加权损失函数。其中:

[0060] (1) 对角对数正则化项:

$$[0061] \quad \mathcal{L}_{dl} = -\frac{\gamma}{n^2} \sum_{i=1}^n \log(|A_{ii}|_{[0, 1]} + \epsilon)$$

[0062] 其中: \mathcal{L}_{dl} 表示对角对数正则化损失, ϵ 表示噪声误差,很小的一个随机参数, A_{ii} 是对角线上的元素, γ 表示自适应因子, A_{ii} 应该大于0并接近于1,因此引入了一个对角对数正则化项来改进预测。

[0063] (2) Kullback-Leibler散度:

[0064] 采用衡量潜在变量分布与单位高斯分布之间差异的Kullback-Leibler散度作为损失函数的一部分,K表示分割的种类数,D表示标准差矩阵,M表示平均值矩阵,Kullback-Leibler散度表达式如下:

$$[0065] \quad \mathcal{L}_{kl} = -\frac{1}{2NK} \sum_{i=1}^N \sum_{j=1}^K (1 + \log(D_{ij})^2 - M_{ij}^2 - (D_{ij})^2)$$

[0066] (3) 自适应多类加权 (ACW) 损失函数:

[0067] 采用自适应多类加权损失函数来解决类的高度不平衡分布。MEAN表示求平均, $\widetilde{\omega}_{ij}$ 适应性广播到像素级系数, p_{ij} 表示像素点, C 表示像素种类数, 自适应多类加权 (ACW) 损失函数表达式如下:

$$[0068] \quad \mathcal{L}_{ACW} = \frac{1}{|Y|} \sum_{i \in Y} \sum_{j \in C} \widetilde{\omega}_{ij} * p_{ij} - \log(\text{MEAN}\{d_j \mid j \in C\})$$

[0069] 其中: \mathcal{L}_{ACW} 表示自适应多类加权损失; Y 包括所有标记像素; d_j 表示骰子系数, 其表达式为:

$$[0070] \quad d_j = \frac{2 \sum_{i \in Y} y_{ij} \widetilde{y}_{ij}}{\sum_{i \in Y} y_{ij} + \sum_{i \in Y} \widetilde{y}_{ij}}$$

[0071] 其中: y_{ij} 表示真实值, 取值范围为 (0, 1); \widetilde{y}_{ij} 表示预测值, 取值范围为属于 (0, 1)。

[0072] 为了细化最终的预测结果, 总损失函数通过将对抗损失与循环一致性损失和几何一致性约束相结合, 采用三种损失函数的总和作为框架中的最终损失函数, 总损失函数可以表述如下:

$$[0073] \quad \mathcal{L} = \mathcal{L}_{dl} + \mathcal{L}_{kl} + \mathcal{L}_{ACW}。$$

[0074] 在其中一个实施例中, 遥感图像语义分割网络的结构如图2所示。步骤104包括: 将训练样本输入到特征图提取网络中, 得到遥感图像的特征图; 将遥感图像的特征图进行图像增强, 得到4张增强特征图; 将4张增强特征图输入到卷积神经网络中, 得到4个自构建邻接矩阵; 将4个自构建邻接矩阵分别输入到4个共享参数的SGA-Net网络得到4个通道相关性特征图; 将4个通道注意力特征图输入到融合预测网络中, 得到遥感图像中每个像素点的预测类别; 根据预测类别、训练样本以及总损失函数, 对遥感图像语义分割网络进行反向训练, 得到训练好的遥感图像语义分割模型。

[0075] 自构建邻接矩阵确保了每张遥感图像的独特性。

[0076] 在其中一个实施例中, SGA-Net网络包括: 图注意力神经网络和通道线性注意力机制; 步骤104还包括: 将第一个自构建邻接矩阵输入到第一个SGA-Net网络的图注意力神经网络中, 得到全局空间相似特征图; 将全局空间相似特征图输入到第一个SGA-Net网络的通道线性注意力机制中, 得到第一个通道相关性特征图; 将第二个自构建邻接矩阵输入到第二个SGA-Net网络中, 得到第二个通道相关性特征图; 将第三个自构建邻接矩阵输入到第三个SGA-Net网络中, 得到第三个通道相关性特征图; 将第4个自构建邻接矩阵输入到第4个SGA-Net网络中, 得到第4个通道相关性特征图。

[0077] 具体的, 图2中遥感图像语义分割网络的特征图提取网络输出遥感图像的特征图 $[C, W, H]$, 其中 C 为通道数, W 为宽, H 为高, 作为优选, $W=H=512$ 。对遥感图像的特征图 $[C, W, H]$ 进行图像增强后, 得到4张增强特征图, 分别为: $X: [C', W', H']$, X_{90} , X_{180} , X_{270} 均是 $[C', W',$

H']。作为优选, $C'=1024,W'=32,H'=32$ 。

[0078] 将4张增强特征图输入到卷积神经网络中,得到4个自构建邻接矩阵 A_0,A_1,A_2,A_3 ,自构建邻接矩阵 A_0,A_1,A_2,A_3 均为 $N\times N$ 阶矩阵,其中 $N=W'\times H'$ 。作为优选, $W'=32,H'=32,N=1024$ 。

[0079] SGA-Net网络中的图注意力神经网络是一个8头的图注意力神经网络。

[0080] 网络中的参数说明: $\vec{h}_1, \vec{h}_2, \vec{h}_3$ 分别表示图像特征 $R^F:F$ 表示特征数,即图像通道数; $\vec{\alpha}_{i1}, \vec{\alpha}_{i2}, \vec{\alpha}_{i3}$ 为 $N\times N$ 阶矩阵,作为优选,为 1024×1024 阶矩阵。 $\vec{h}'_1, \vec{h}'_2, \vec{h}'_3$ 为 $N\times 8$ 阶矩阵,其中8表示8头注意力的图注意力机制;图注意力神经网络输出的全局空间相似特征图 h 为 $N\times T$ 阶矩阵,其中: T 表示类别数,作为优选 h 为 1024×6 阶矩阵。

[0081] 在其中一个实施例中,步骤104还包括:将全局空间相似特征图输入到通道线性注意力机制中,分别对全局空间相似特征图进行3次变形操作,得到第一变形特征、第二变形特征以及第三变形特征;对全局空间相似特征图进行变形操作,并对变形得到的结果进行转置,得到转置变形特征;将第三变形特征与转置变形特征相乘,并将得到的特征进行变形操作,得到特征矩阵;将特征矩阵与第二变形特征进行相乘,并将得到的特征进行变形操作后与第一变形特征相加,得到第一个通道相关性特征图。

[0082] SGA-Net网络中的通道线性注意力机制中的参数说明: Q,K 是全局空间相似特征图进行变形操作后得到的第二变形特征和第三变形特征, Q,K 以及第一变形特征均为 $T\times N$ 阶矩阵,其中 N 表示低等特征图的高和宽的积, T 表示种类数; V 是全局空间相似特征图进行变形操作后再进行转置得到的转置变形矩阵, V 为 $N\times T$ 阶矩阵,并将 V 变换为 $T\times N$ 阶矩阵;将 $V\times K$ 得到的结果进行变形得到特征矩阵 $[T,P,P]$,其中: $P\times P=N,P=W'=H'$;通道线性注意力机制输出通道相关性特征图是 $[T,P,P]$ 。

[0083] 在其中一个实施例中,图注意力神经网络是8头图注意力神经网络。

[0084] 自构建邻接矩阵被输入到图注意力神经网络中,因为遥感图像是RGB三通道,因此采用8头的图注意力神经网络,可以增强网络的稳定性。

[0085] 在其中一个实施例中,步骤104还包括:将遥感图像的特征图旋转 $0^\circ, 90^\circ, 180^\circ$ 以及 270° ,得到4张增强特征图。

[0086] 应该理解的是,虽然图1的流程图中的各个步骤按照箭头的指示依次显示,但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明,这些步骤的执行并没有严格的顺序限制,这些步骤可以以其它的顺序执行。而且,图1中的至少一部分步骤可以包括多个子步骤或者多个阶段,这些子步骤或者阶段并不必然是在同一时刻执行完成,而是可以在不同的时刻执行,这些子步骤或者阶段的执行顺序也不必然是依次进行,而是可以与其它步骤或者其它步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

[0087] 在一个验证性实施例中,采用公开数据集对本方法进行验证。

[0088] (1)数据集

[0089] 数据集包括两个公开数据集,分别为:ISPRS Potsdam公开数据集,一共38张遥感图像;分辨率:5cm;训练集:22张;验证集:2张;测试集:14张;大小:6000×6000;ISPRS Vaihingen(公开数据集)一共33张遥感图像;分辨率:9cm;训练集:14张;验证集:2张;测试集:17张;大小:不定;

[0090] (2)评价指标

[0091] 评价指标包括:mIoU(Mean Intersection over Union) (均交并比)、F1分数、Acc (Pixel Accuracy(%))准确性。

[0092] mIoU:计算真实值和预测值两个集合的交集和并集之比。

[0093] 语义分割说到底还是一个分类任务,既然是分类任务,预测的结果往往就是四种情况:(1) true positive (TP):真实值;(2) false positive (FP):假正值;(3) true negative(TN):真负值;(4) false negative (FN):假负值。

$$[0094] \quad mIoU = \frac{1}{N} \sum_{k=1}^N \frac{TP_k}{TP_k + FP_k + FN_k}$$

[0095] mIoU越大越好,代表遥感图像语义分割越准确其中。

$$[0096] \quad F1 = 2 \times \frac{precision \times recall}{precision + recall}$$

$$[0097] \quad recall = \frac{TP}{TP + FN}$$

$$[0098] \quad precision = \frac{TP}{TP + FP}$$

[0099] F1分数越高,表示分类模型越稳健。

$$[0100] \quad Acc = \frac{\sum_{k=1}^N (TP_k + TN_k)}{\sum_{k=1}^N (TP_k + FP_k + TN_k + FN_k)}$$

[0101] 表示分类准确像素的占比,Acc越大,表示分类越准确。

[0102] (3)实验环境

[0103] 使用Python 3.6和PyTorch框架进行实验。由于网络的复杂性和损失函数的计算量大,我们使用NVIDIA Tesla V100 GPU和NVIDIA的CUDA 10.0 API模型来训练模型。

[0104] (4)实验结果

[0105] 为了验证遥感图像语义分割的有效性,我们将该方法与其他五种方法在两个数据集上进行比较,具体数据如表1、表2所示。。

[0106] 表1不同方法使用Potsdam数据集的得分

Method	Road surf	Buildings	Low veg.	Trees	Cars	Mean F1	Acc	mIoU
MSCG-Net	<u>0.907</u>	<u>0.926</u>	0.851	0.872	0.911	0.893	0.959	0.807
DANet	<u>0.907</u>	0.922	0.853	0.868	0.919	0.894	0.959	0.807
Deeplab V3	0.905	0.924	0.850	0.870	<u>0.939</u>	0.897	0.958	0.806
DUNet	<u>0.907</u>	0.925	0.853	0.869	<u>0.935</u>	0.898	0.959	<u>0.808</u>
DDCM	0.901	0.924	<u>0.871</u>	<u>0.890</u>	0.932	<u>0.904</u>	<u>0.961</u>	<u>0.808</u>
SGA-Net	0.927	0.958	0.886	0.896	0.968	0.927	0.964	0.832

[0108] 表2不同方法使用Vaihingen数据集的得分

	Method	Road surf	Buildings	Low veg.	Trees	Cars	Mean F1	Acc	mIoU
[0109]	MSCG-Net	0.906	0.924	0.816	<u>0.887</u>	0.820	0.870	0.955	0.796
	DANet	0.905	0.934	0.833	<u>0.887</u>	0.761	0.859	0.955	0.797
	Deeplab V3	0.911	0.927	0.819	0.886	0.818	0.872	0.956	0.800
	DUNet	0.910	0.927	0.817	<u>0.887</u>	0.843	0.877	0.955	0.801
	DDCM	<u>0.927</u>	<u>0.953</u>	0.833	0.890	<u>0.883</u>	0.898	0.963	0.828
	SGA-Net	0.932	0.955	0.826	0.884	0.928	0.905	0.965	<u>0.826</u>

[0110] 从两个数据集中随机选取几个生成图像的样本。显然,SGA-Net可以提高两个数据集的分割性能,性能优于之前的方法。

[0111] 在一个实施例中,如图3所示,提供了一种遥感图像语义分割装置,包括:遥感图像获取模块、遥感图像语义分割网络构建模块、遥感图像语义分割网络训练模块以及遥感图像像素点类别确定模块,其中:

[0112] 遥感图像获取模块,用于获取RGB三通道的遥感图像,并将遥感图像作为训练样本。

[0113] 遥感图像语义分割网络构建模块,用于构建遥感图像语义分割网络;遥感图像语义分割网络包括:特征图提取网络、卷积神经网络、4个共享参数的SGA-Net网络以及融合预测模块;特征提取网络用于提取遥感图像的特征图;卷积神经网络用于提取图像增强后的遥感图像的特征图的自构建邻接矩阵;SGA-Net网络采用图注意力神经网络挖掘遥感图像各个像素点之间的空间相似度,并采用通道线性注意力机制捕捉神经网络通道输出之间的相关性;融合预测模块用于将4个SGA-Net网络输出的特征进行融合预测,得到遥感图像中每个像素点的类别。

[0114] 遥感图像语义分割网络训练模块,用于根据训练样本对遥感图像语义分割网络进行训练,得到训练好的遥感图像语义分割模型。

[0115] 遥感图像像素点类别确定模块,用于获取待测RGB三通道的遥感图像,并输入到遥感图像语义分割模型中,得到遥感图像中每个像素点的类别。

[0116] 在其中一个实施例中,遥感图像语义分割网络训练模块前还包括总损失函数构建模块,用于构建总损失函数,总损失函数包括:对角对数正则化项、Kullback-Leibler散度以及自适应多类加权损失函数。

[0117] 在其中一个实施例中,遥感图像语义分割网络训练模块,还用于将训练样本输入到特征图提取网络中,得到遥感图像的特征图;将遥感图像的特征图进行图像增强,得到4张增强特征图;将4张增强特征图输入到卷积神经网络中,得到4个自构建邻接矩阵;将4个自构建邻接矩阵分别输入到4个共享参数的SGA-Net网络得到4个通道相关性特征图;将4个通道注意力特征图输入到融合预测网络中,得到遥感图像中每个像素点的预测类别;根据预测类别、训练样本以及总损失函数,对遥感图像语义分割网络进行反向训练,得到训练好的遥感图像语义分割模型。

[0118] 在其中一个实施例中,SGA-Net网络包括:图注意力神经网络和通道线性注意力机制;遥感图像语义分割网络训练模块,还用于将第一个自构建邻接矩阵输入到第一个SGA-Net网络的图注意力神经网络中,得到全局空间相似特征图;将全局空间相似特征图输入到第一个SGA-Net网络的通道线性注意力机制中,得到第一个通道相关性特征图;将第二个自构建邻接矩阵输入到第二个SGA-Net网络中,得到第二个通道相关性特征图;将第三个自构建邻接矩阵输入到第三个SGA-Net网络中,得到第三个通道相关性特征图;将第4个自构建邻接矩阵输入到第4个SGA-Net网络中,得到第4个通道相关性特征图。

[0119] 在其中一个实施例中,遥感图像语义分割网络训练模块,还用于将全局空间相似特征图输入到通道线性注意力机制中,分别对全局空间相似特征图进行3次变形操作,得到第一变形特征、第二变形特征以及第三变形特征;对全局空间相似特征图进行变形操作,并对变形得到的结果进行转置,得到转置变形特征;将第三变形特征与转置变形特征相乘,并将得到的特征进行变形操作,得到特征矩阵;将特征矩阵与第二变形特征进行相乘,并将得到的特征进行变形操作后与第一变形特征相加,得到第一个通道相关性特征图。

[0120] 在其中一个实施例中,图注意力神经网络是8头图注意力神经网络。

[0121] 在其中一个实施例中,遥感图像语义分割网络训练模块,还用于将遥感图像的特征图旋转 0° 、 90° 、 180° 以及 270° ,得到4张增强特征图。

[0122] 关于遥感图像语义分割装置的具体限定可以参见上文中对于遥感图像语义分割方法的限定,在此不再赘述。上述遥感图像语义分割装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中,也可以以软件形式存储于计算机设备中的存储器中,以便于处理器调用执行以上各个模块对应的操作。

[0123] 在一个实施例中,提供了一种计算机设备,该计算机设备可以是终端,其内部结构图可以如图4所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口、显示屏和输入装置。其中,该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统和计算机程序。该内存储器为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机程序被处理器执行时以实现一种遥感图像语义分割方法。该计算机设备的显示屏可以是液晶显示屏或者电子墨水显示屏,该计算机设备的输入装置可以是显示屏上覆盖的触摸层,也可以是计算机设备外壳上设置的按键、轨迹球或触控板,还可以是外接的键盘、触控板或鼠标等。

[0124] 本领域技术人员可以理解,图4中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的计算机设备的限定,具体的计算机设备可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。

[0125] 在一个实施例中,提供了一种计算机设备,包括存储器和处理器,该存储器存储有计算机程序,该处理器执行计算机程序时实现上述方法实施例中的步骤。

[0126] 在一个实施例中,提供了一种计算机可读存储介质,其上存储有计算机程序,计算机程序被处理器执行时实现上述方法实施例中的步骤。

[0127] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一非易失性计算机可读存储介质中,该计算机程序在执行时,可包括如上述各方法的实施例的流程。其中,本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用,均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲存储器。作为说明而非局限,RAM以多种形式可得,诸如静态RAM(SRAM)、动态RAM(DRAM)、同步DRAM(SDRAM)、双数据率SDRAM(DDRSDRAM)、增强型SDRAM(ESDRAM)、同步链路(Synchlink)DRAM(SLDRAM)、存储器总线(Rambus)直接RAM

(RDRAM)、直接存储器总线动态RAM(DRDRAM)、以及存储器总线动态RAM(RDRAM)等。

[0128] 以上实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

[0129] 以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但不能因此而理解为对发明专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本申请构思的前提下,还可以做出若干变形和改进,这些都属于本申请的保护范围。因此,本申请专利的保护范围应以所附权利要求为准。

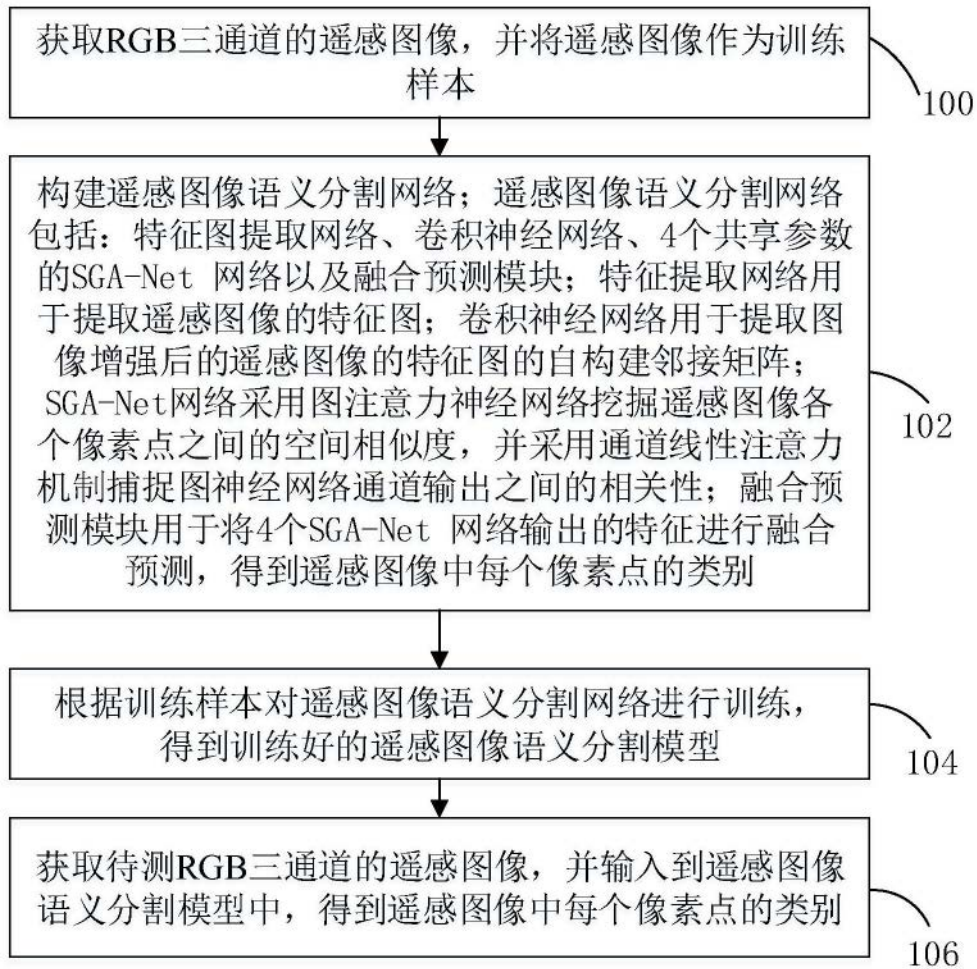


图1

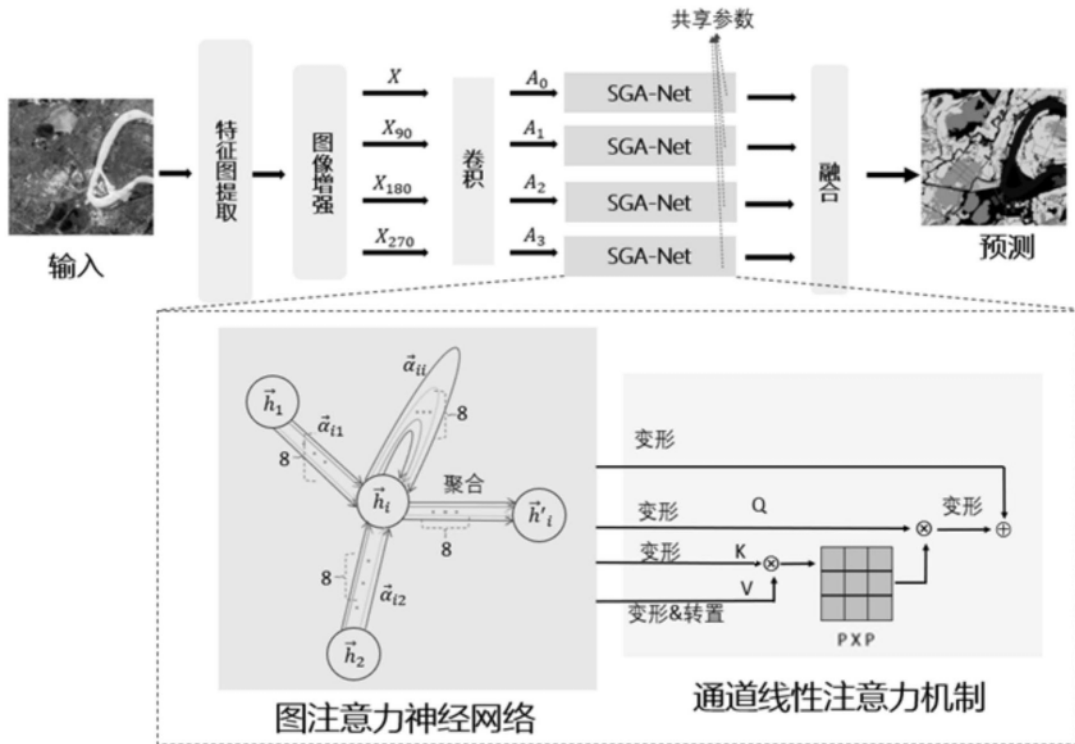


图2



图3

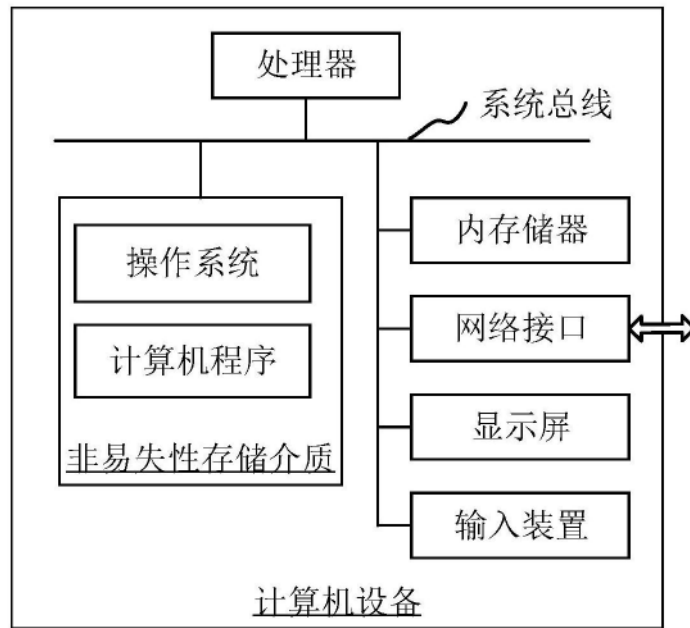


图4