



(12) 发明专利

(10) 授权公告号 CN 118918331 B

(45) 授权公告日 2025. 01. 28

(21) 申请号 202411397518.3

(22) 申请日 2024.10.09

(65) 同一申请的已公布的文献号

申请公布号 CN 118918331 A

(43) 申请公布日 2024.11.08

(73) 专利权人 三峡智能工程有限公司

地址 430000 湖北省武汉市东湖新技术开发区九峰街道高新大道789号全球公共采购交易服务总部基地办公部分二期4号楼9层1室(自贸区武汉片区)

(72) 发明人 王海羽 金和平 刘洁 罗惠恒

杨磊 胡金艳 李林泽 安日辉

涂晶 徐章智

(74) 专利代理机构 武汉开元知识产权代理有限公司

公司 42104

专利代理师 徐锐 陈家安

(51) Int.Cl.

G06V 10/26 (2022.01)

G06V 10/44 (2022.01)

G06V 10/75 (2022.01)

G06V 10/82 (2022.01)

G06V 20/17 (2022.01)

G06N 3/0475 (2023.01)

G06N 3/045 (2023.01)

G06N 3/094 (2023.01)

(56) 对比文件

CN 114359360 A, 2022.04.15

CN 115049841 A, 2022.09.13

审查员 徐盛辉

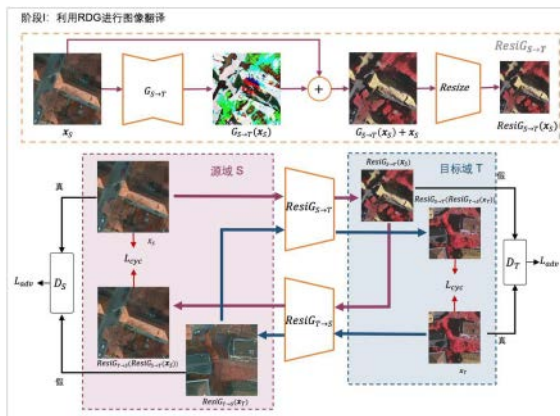
权利要求书1页 说明书8页 附图1页

(54) 发明名称

一种遥感图像深度学习模型的构建方法、记录媒体及系统

(57) 摘要

本发明属于图像数据处理技术领域,特别涉及一种遥感图像深度学习模型的构建方法,本发明在基于循环一致性的生成对抗网络模型基础上,面向无人机遥感影像的非监督图像转换和跨领域语义分割提出了一种包含残差-重采样的对偶生成式对抗网络模型的构建方法,通过该方法构建的模型使用重采样模块来充分利用无人机影像的尺度信息,并通过残差连接,将生成器的功能从生成新图像变为生成残差信息,从而增强语义信息在生成过程中的保留程度,并在自适应框架的末端加入一个判别器,用于优化输出空间的与语义分割结果。本发明还提供一种存储有该方法程序的可读记录媒体及包含该媒体的系统,通过处理电路可以调用程序,执行上述方法。



1.一种遥感图像深度学习模型的构建方法,其特征在于包括以下步骤:

S1.构建双向生成式对抗神经网络,所述双向生成式对抗神经网络包括:

源域至目标域重采样残差的生成器 $\text{ResiG}_{S \rightarrow T}$,用于执行从源域S到目标域T的图像转换,并通过重采样技术调整图像分辨率与目标域一致;

目标域至源域重采样残差的生成器 $\text{ResiG}_{T \rightarrow S}$,用于执行目标域T到源域S的图像转换,并通过重采样技术调整图像分辨率与源域一致;

判别器 D_S ,用于通过对抗训练判断图像是来自源域S,还是通过 $\text{ResiG}_{T \rightarrow S}$ 生成,判别器 D_T ,用于通过对抗训练判断图像是来自目标域T,还是通过 $\text{ResiG}_{S \rightarrow T}$ 生成;

S2.利用双向生成式对抗神经网络获取目标域化的源域数据 $X_{S \rightarrow T}$ 及对应的标记数据 Y_S ,设置有监督的训练一个语义分割模型 f_T 用于无标签目标域数据 X_T 的预测;在语义分割模型 f_T 的输出空间设置一个能通过对抗训练判别输出空间的分割结果所属领域的判别器 D_{out} ,完成遥感图像深度学习模型的构建;

所述对抗训练通过迭代计算判别器损失 L_{adv} ,直至 L_{adv} 减小至设定的阈值内,完成训练;

S1步骤中的对抗训练还包括以下步骤:

计算同域原始图像和经生成器转至他域后再从他域转回的图像间的重建损失 L_{cyc} ;通过交替迭代计算 L_{cyc} 与 L_{adv} ,并与预设阈值比较,训练生成器与判别器。

2.根据权利要求1所述的一种遥感图像深度学习模型的构建方法,其特征在于,所述调整图像分辨率与目标域一致的方法是通过重采样函数,将源域调整到目标域数据类似的空间尺度的过程,选择双线性内插函数作为重采样函数。

3.根据权利要求2所述的一种遥感图像深度学习模型的构建方法,其特征在于,生成器是一个包括编码器和解码器的全卷积网络,编码器和解码器之间存在跳跃连接,编码器作为下采样层,设定卷积核大小为4,填充为1,步长为2,各卷积层的卷积核深度分别为64、128、256、512、512、512、512,其中除了第一层和最后一层,各个卷积层经过了实例归一化,包含负斜率为0.2的Leaky ReLU激活函数;解码器作为上采样层,采用了与编码器相同的卷积核深度、大小、步长和填充参数,编码器和解码器中卷积核深度超过256的层中还采用了概率为0.5的移出层。

4.根据权利要求3所述的一种遥感图像深度学习模型的构建方法,其特征在于,判别器采用全卷积网络,各层的卷积核深度为64、128、256、512、512、1,除最后一层以外,其它层均包含批归一化层和负斜率为0.2的Leaky ReLU激活函数。

5.一种非暂态可读记录媒体,用以存储包含多个指令的一个或多个程序,其特征在于,当执行指令时,将致使处理电路执行权利要求1-4中任一项所述的一种遥感图像深度学习模型的构建方法。

6.一种遥感图像深度学习模型的构建系统,包括处理电路及与其电性耦接的存储器,其特征在于,所述存储器配置储存至少一程序,所述程序包含多个指令,所述处理电路运行所述程序,能执行权利要求1-4中任一项所述的一种遥感图像深度学习模型的构建方法。

一种遥感图像深度学习模型的构建方法、记录媒体及系统

技术领域

[0001] 本发明属于图像数据处理技术领域,公开了一种遥感图像深度学习模型的构建方法、记录媒体及系统。

背景技术

[0002] 语义分割技术在水体区域提取、水体周围环境要素识别等任务中起着重要作用。然而,由于深度学习模型自身的局限性,当在标注好的训练集上训练好的模型部署到真实环境中时,由于真实环境和训练集影像在成像设备、时间、区域等方面的差异,模型的性能往往会出现下降,域适应方法则是为了提升模型在这种情况下表现。现有的领域自适应(简称“域适应”,下同)方法存在两个主要问题:一是忽略了无人机影像翻译“真实1-真实2”的特质(传统计算机视觉图像翻译为GTA5数据集到Cityscape的虚拟-真实翻译);二是忽视了不同分辨率对域适应的影响,这就导致深度学习的结果与无人机采集图像的真实期望结果存在偏差。

[0003] 基于对抗学习的自适应方法的核心在于通过训练判别器识别输入的信息是否来自于特定的数据域而让模型学习到这种数据的分布差异。判别器可以在整个深度学习框架的不同阶段介入,从而实现两个领域的对齐或是让网络的特征提取器以及解码器学习到两个领域的分布差异,提高在目标域数据上的表现,但此类技术有两个主要的局限性。首先,对齐边缘分布并不强制执行任何语义一致性,例如,水体的目标特征可能被映射到建筑的源特征。其次,在深层表征的较高层次上的对齐可能无法模拟对最终视觉任务至关重要的低层次外观差异。

[0004] 就遥感图像语义分割的领域自适应而言,利用基于生成式对抗网络的无监督领域自适应方法使得高分遥感图像的跨域语义分割的性能有了很大的提高。而用于遥感图像翻译的生成式对抗网络结构大多直接引用计算机视觉领域的模型,如CycleGAN 和DualGAN 等。但遥感图像与计算机视觉领域任务中日常图像之间存在着较大的差异,因此直接使用对应的模型可能带来以下问题:

[0005] 首先,目前现有的无监督图像翻译算法无法解决遥感图像跨领域的空间尺度差异,而尺度是对地观测图像的重要属性之一。由于航天或航空对地观测的光学传感器通常具有固定焦距,传感器视角和高度的变化会直接影响成像的分辨率,所以通常使用固定相机角度在固定高度上进行成像作业。相同的地物目标在空间分辨率不同的遥感图像中所占据的像素数量是不同的。而在计算机视觉领域,以用于自动驾驶的车载相机为例,在运动成像过程中,不同的物体根据距离相机位置远近尺度是不断变化的,在不同的街景数据集中,目标对象都是以多尺度的状态而存在,因此不存在遥感图像中的尺度效应问题。所以在进行遥感图像跨领域图像转换时,空间尺度差异问题需要进行额外的网络模块进行对齐,否则会导致特定尺度敏感的地物类别的分割精度的下降。考虑到领域数据之间的空间尺度差异导致的尺度效应本身就是数据分布差异的来源之一,因此显式的进行空间尺度转换是减少分布差异的重要途径。

[0006] 其次,计算机视觉进行图像转换的领域除了现实世界的图像外、还包括绘画、游戏等非现实图像数据,因此生成器和判别器在设计时并不会针对局部的语义信息进行保留或增强,而只是让生成数据看起来来自对应的领域。这种网络结构可能会造成关键语义信息的丢失。而遥感图像的跨领域自适应均为现实世界信息,具备真实地理意义。图像翻译过程中生成的数据很可能会给某个地物对象带来结构信息的变化,从而影响后续的语义分割的任务。

发明内容

[0007] 针对上述问题,本发明提供一种遥感图像深度学习模型的构建方法,具体方案包括以下步骤:

[0008] S1. 构建双向生成式对抗神经网络,所述双向生成式对抗神经网络包括:

[0009] 源域至目标域重采样残差的生成器 $ResiG_{S \rightarrow T}$,用于执行从源域 S 到目标域 T 的图像转换,并通过重采样技术调整图像分辨率与目标域一致;

[0010] 目标域至源域重采样残差的生成器 $ResiG_{T \rightarrow S}$,用于执行目标域 T 到源域 S 的图像转换,并通过重采样技术调整图像分辨率与源域一致;

[0011] 判别器 D_S ,用于通过对抗训练判断图像是来自源域 S,还是通过 $ResiG_{T \rightarrow S}$ 生成,判别器 D_T ,用于通过对抗训练判断图像是来自目标域 T,还是通过 $ResiG_{S \rightarrow T}$ 生成;

[0012] S2. 利用双向生成式对抗神经网络获取目标域化的源域数据 $X_{S \rightarrow T}$ 及对应的标记数据 Y_S ,设置有监督的训练一个语义分割模型 f_T 用于无标签目标域数据 X_T 的预测;在语义分割模型 f_T 的输出空间设置一个能通过对抗训练判别输出空间的分割结果所属领域的判别器 D_{out} ,完成遥感图像深度学习模型的构建。

[0013] 优选的,所述对抗训练通过迭代计算判别器损失 L_{adv} ,直至 L_{adv} 减小至设定的阈值内,完成训练。

[0014] 优选的,还包括以下步骤:

[0015] 计算同域原始图像和经生成器转至他域后再从他域转回的图像间的重建损失 L_{cyc} ;通过交替迭代计算 L_{cyc} 与 L_{adv} ,并与预设阈值比较,训练生成器与判别器。

[0016] 优选的,所述调整图像分辨率与目标域一致的方法是通过重采样函数,将源域调整到目标域数据类似的空间尺度的过程,选择双线性内插函数作为重采样函数。

[0017] 优选的,生成器是一个包括编码器和解码器的全卷积网络,编码器和解码器之间存在跳跃连接,编码器作为下采样层,设定卷积核大小为4,填充为1,步长为2,各卷积层的卷积核深度分别为64、128、256、512、512、512、512,其中除了第一层和最后一层,各个卷积层经过了实例归一化,包含负斜率为0.2的Leaky ReLU激活函数;解码器作为上采样层,采

用了与编码器相同的卷积核深度、大小、步长和填充参数,编码器和解码器中卷积核深度超过256的层中还采用了概率为0.5的移出层。

[0018] 优选的,判别器采用全卷积网络,各层的卷积核深度为64、128、256、512、512、1,除最后一层以外,其它层均包含批归一化层和负斜率为0.2的Leaky ReLU激活函数。

[0019] 本发明的另一方案在于提供一种可读记录媒体,用以存储包含多个指令的一个或多个程序,当运行程序时,将致使处理电路执行上述遥感图像深度学习模型的构建方法。

[0020] 本发明的又一方案在于提供一种遥感图像深度学习模型的构建系统,包括处理电路及与其电性耦接的存储器,所述存储器配置储存至少一程序,所述程序包含多个指令,所述处理电路运行所述程序,能执行上述遥感图像深度学习模型的构建方法。

[0021] 相对于现有技术,本发明产生以下有益效果:

[0022] 通过本发明提出方法构建的遥感图像深度学习模型,旨在从有标注的数据集向无标注的数据集学习遥感图像的语义分割模型,以最小化像素级的领域差距。考虑到无人机影像中尺度不一的物体的尺度差异,使用了一个重采样模块,提高了尺度不一的类别的准确性。

[0023] 考虑到无人机影像实对实翻译的特点,同时使用了残差连接。与重采样模块结合,残差连接结构实现跨域数据分布匹配性能方面发挥着关键作用。残差连接通过恒等连接保留了原始数据的语义信息,避免了结构信息的修改。无人机影像中包含的是对地观测的真实场景,因此领域之间的转换同样是“现实”到“现实”的,其中所有像元都具有地理意义。同时,模型应尽可能少的修改可能会影响后续分割任务的真实场景地物结构信息。因此在模型中专门设计了残差连接结构,尽可能地保留了原始结构信息,并侧重于其他基础特征的转换,例如颜色、阴影等。因此,残差连接提高了有效提升后续的分割性能,更适合无人机影像域之间的互相转换。此外,残差连接还可以稳定模型的训练过程。传统对偶生成对抗网络的生成器的作用是从输入图像中生成一个目标风格化的图像,而本方法的生成器的作用是从输入图像中生成一个残差项,将传统对偶生成对抗网络中的生成器从生成具有目标样式的新图像的负担中解放出来,并在很大程度上保留了输入图像的信息,从而稳定了训练过程。

[0024] 本发明同时结合输出空间的适应性,进一步提高模型的性能,整个对抗学习训练框架在基准上达到了较高的精度表现,同时具有结构简单、稳定性好的优点。

附图说明

[0025] 图1为本发明实施例中遥感图像生成式残差-重采样对抗网络模型的数据流转示意图;

[0026] 图2为本发明实施例中借监督语义分割子模型与输出空间对抗学习示意图。

具体实施方式

[0027] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行描述,所描述的实施例是本发明一部分实施例,而不是全部实施例。基于本发明中的实施例,本领域技术人员在没有做出创新劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0028] 下面结合附图和实施例对本发明做进一步详细说明。

[0029] 如图1所示,一种遥感图像深度学习模型的构建方法,该模型的构建包含两个阶段。阶段I使用生成式空间尺度转换对抗网络 (ResiDualGAN),以便源域到目标域进行无监督图像转换;阶段II则通过阶段I生成的目标域化 (target-stylized) 的源域数据和对应的标签数据对语义分割模型 f_T 进行有监督的训练,建立类目标域数据分布与标注数据之间的映射关系。为了优化目标域的分割结果,同时增加对抗学习在本框架中的应用,在分割模型的末端输出空间增加了一个输出空间对抗学习判别网络模块,以提高跨领域语义分割模型的性能。下面对两个阶段的训练过程进行详细介绍。

[0030] 一、生成式残差-重采样对抗网络

[0031] 如图1所示,阶段I的主要目标是通过生成式对抗网络在输入空间进行像素级数据分布对齐,生成目标域化的源域数据,其表现为将 X_S 转换为类似于 X_T 的数据分布。为此本发明设计了一种生成式残差-重采样对抗网络用于输入空间的图像数据分布特征匹配,命名为ResiDualGAN。与常规的生成对抗网络类似,ResiDualGAN由对偶的生成器 (generators) 和判别器 (discriminators) 组成。不同点在于ResiDualGAN中的重采样残差生成器 (ResiGenerator) 能够在生成图像的同时进行重采样。判别器则用于判断输入图像是否是ResiGenerator生成的图像,进行对抗学习。具体来说,ResiDualGAN 包含两个重采样残差生成器: $ResiG_{S \rightarrow T}$ 和 $ResiG_{T \rightarrow S}$, 以及两个判别器: D_S 和 D_T 。其中 $ResiG_{S \rightarrow T}$ 用于执行从源域 S 到目标域 T 的图像转换与重采样, $ResiG_{T \rightarrow S}$ 执行目标域 T 到源域 S 的图像转换与重采样。而判别器 D_S 用于判断图像是来自源域 S 还是来自 $ResiG_{T \rightarrow S}$, 判别器 D_T 用于判断图像是来自目标域 T 还是来自 $ResiG_{S \rightarrow T}$ 生成。ResiDualGAN进行图像风格转换的原理如下:随机输入源域数据 X_S 与目标域数据 X_T , 源域输入图像 X_S 通过尺度转换生成器 $ResiG_{S \rightarrow T}$ 后得到转换到目标域,得到类目标域 (target-stylized) 的源域数据 $ResiG_{S \rightarrow T}(X_S)$ 。 $ResiG_{S \rightarrow T}(X_S)$ 与 X_T 被送入目标域判别器 D_T , D_T 试图区分判别器网络的输入是来自源域生成还是目标域本身,计算判别损失 L_{adv} 从而进行对抗学习。 $ResiG_{S \rightarrow T}(X_S)$ 再通过反向的生成器 $ResiG_{T \rightarrow S}$ 转换回源域,得到重建后的图像 $ResiG_{T \rightarrow S}(ResiG_{S \rightarrow T}(X_S))$ 。基于循环一致性原则,理想情况下,通过正向与反向生成器的源域输入图像 X_S 应与与其重建的图像 $ResiG_{T \rightarrow S}(ResiG_{S \rightarrow T}(X_S))$ 保持一致。因此可以用两者之间的差异计算一致性损失(或称为重建损失) L_{cyc} 。目标域输入数据 X_T 通过网络的过程是上述输入 X_S 的逆过程,在此不再赘述。通过交替计算重建损失与判别器损失,训练两个生成器与判别器,进行双向的

生成对抗学习。

[0032] 阶段I的建模过程：

[0033] (1) 重采样残差生成器子模型构建

[0034] 源域到目标域的转换过程 $\mathbb{R}^{H_S \times W_S \times C} \rightarrow \mathbb{R}^{H_T \times W_T \times C}$ 如图2所示。该过程通过重采样残差生成器 (ResiGenerator) 完成, 将源域数据转换为类目标域数据的同时完成重采样。ResiGenerator包含一个基于全卷积神经网络的生成器 $G_{S \rightarrow T}$ 和一个重采样模块 $Resize_{S \rightarrow T}$ 。生成器 $G_{S \rightarrow T}: \mathbb{R}^{H_S \times W_S \times C} \rightarrow \mathbb{R}^{H_T \times W_T \times C}$ 从随机采样选择的源域数据中生成一个残差项, 重采样模块 $Resize_{S \rightarrow T}: \mathbb{R}^{H_S \times W_S \times C} \rightarrow \mathbb{R}^{H_T \times W_T \times C}$ 是一个重采样函数, 将源域调整到目标域数据类似的空间尺度。根据消融实验, 我们选择了双线性内插函数作为该模块使用的函数。重采样残差生成器将 X_S 转化为目标域化的 $X_{S \rightarrow T}$ 的过程如下所示:

$$[0035] \quad X_{S \rightarrow T} = ResiG_{S \rightarrow T}(X_S) = Resize_{S \rightarrow T}(G_{S \rightarrow T}(X_S) + X_S)$$

[0036] 公式省略了上述过程中涉及的随机噪声。根据空间分辨率尺度一致性原则, 在重采样之后, $X_{S \rightarrow T}$ 的空间分辨率与 X_T 相一致。 $ResiG_{T \rightarrow S}$ 的结构与 $ResiG_{S \rightarrow T}$ 类似, 唯一的区别是重采样模块 $Resize_{T \rightarrow S}: \mathbb{R}^{H_T \times W_T \times B} \rightarrow \mathbb{R}^{H_S \times W_S \times B}$ 是 $Resize_{S \rightarrow T}$ 的反向过程, 将大小为 $H_T \times W_T \times B$ 的图像重采样为大小为 $H_S \times W_S \times B$ 的影像。因此, $ResiG_{T \rightarrow S}(X_{S \rightarrow T})$ 的大小和分辨率应与源域的标记数据 X_S 相同。

[0037] (2) 子模型网络结构设置

[0038] 重采样残差生成器结构由一个生成器和重采样模块组成。生成器是一个包括编码器和解码器的全卷积网络, 编码器和解码器之间存在跳跃连接 (skip connections)。对于编码器 (下采样层), 设定卷积核大小为4, 填充为1, 步长为2。各卷积层的卷积核深度分别为64、128、256、512、512、512、512。其中除了第一层和最后一层, 各个卷积层经过了实例归一化 (instance normalization) 和负斜率为0.2的Leaky ReLU激活函数。解码器 (上采样层) 采用了与编码器相同的卷积核深度、大小、步长和填充等参数, 编码器和解码器中卷积核深度超过256的层中还采用了概率为0.5的Dropout层。重采样模块则使用双线性差值实现。ResiDualGAN中的判别器也是通过全卷积网络实现的。各层的卷积核深度为64、128、256、512、512、1。除最后一层以外, 其它层均包含批归一化层 (batch normalization) 和负斜率为0.2的Leaky ReLU激活函数。由于Wasserstein GAN (WGAN) 的限制, 最后一层的输出不经过任何激活函数。

[0039] (3) 子模型目标优化函数

[0040] 重采样残差生成器生成图像用于训练判别器, 而判别器试图来判定生成图像的来, 从而可以得到对抗学习的损失函数 $\mathcal{L}_{adv}$ 。本发明使用WGAN的损失用于对抗性损失。WGAN通过Earth-Mover (EM) 距离作为数据分布差异的度量标准, 解决了两个数据分布完全没有交集时距离为0的问题, 从而避免梯度消失的问题。与传统的生成式对抗网络使用

Sigmoid输出作为判别器的最终输出不同,ResiDualGAN 移除了Sigmoid层作为最终输出。此外,使用梯度罚项用于稳定训练过程,避免WGAN中权重剪切 (weight clipping) 操作,以实现Lipschitz约束。为简化起见,梯度罚项没有体现在公式中。综上,源域 S 到目标域 T 的对抗性损失 L_{adv} 如下所示:

$$[0041] \quad \mathcal{L}_{adv}^{S \rightarrow T} = \mathbb{E}_{x_T \sim X_T} (D_T(x_T)) - \mathbb{E}_{x_S \sim X_S} (D_T(\text{ResiG}_{S \rightarrow T}(x_S)))$$

[0042] $\text{ResiG}_{T \rightarrow S}$ 将 $X_{S \rightarrow T}$ 重构为源域的风格,以执行对偶学习的循环一致性。理想情况下, $\text{ResiG}_{T \rightarrow S}(X_{S \rightarrow T})$ 应该与 X_S 完全相同。但实际会存在差异,这种差异被用于计算重建损失函数。重建损失函数 $\mathcal{L}_{cyc}$ 可以通过 L1 正则项来衡量:

$$[0043] \quad \mathcal{L}_{cyc} = \mathbb{E}_{x_S \sim X_S} (\| \text{ResiG}_{T \rightarrow S}(\text{ResiG}_{S \rightarrow T}(x_S)) - x_S \|_1) + \mathbb{E}_{x_T \sim X_T} (\| \text{ResiG}_{S \rightarrow T}(\text{ResiG}_{T \rightarrow S}(x_T)) - x_T \|_1)$$

[0044] 考虑到对抗性损失和重建损失,ResiDualGAN的总损失可以写成:

$$[0045] \quad \mathcal{L}_{\text{ResiG}} = \lambda_{cyc} \mathcal{L}_{cyc} + \lambda_{adv} (\mathcal{L}_{adv}^{S \rightarrow T} + \mathcal{L}_{adv}^{T \rightarrow S})$$

[0046] 其中 λ_{cyc} 和 λ_{adv} 是对应于重建损失和判别损失的超参数。在ResiDualGAN的训练中,判别器试图最大化 $\mathcal{L}_{\text{ResiG}}$,而ResiGenerator试图最小化,这可以写成以下最小-最大准则:

$$[0047] \quad \max_{D_S} \min_{\substack{\text{ResiG}_{S \rightarrow T} \\ D_T}} \mathcal{L}_{\text{ResiG}}$$

[0048] 在阶段I的训练结束后,可以得到一个重采样残差生成器 $\text{ResiG}_{S \rightarrow T}$ 来执行空间分辨率尺度一致性原则中的转换任务,生成 $X_{S \rightarrow T}$ 用于阶段II的训练。

[0049] 二、监督语义分割子模型与输出空间对抗学习

[0050] 阶段II的目标是利用ResiDualGAN生成的目标域化的源域数据 $X_{S \rightarrow T}$ 及对应的标记数据 Y_S ,有监督的训练一个语义分割模型 f_T 用于无标签目标域数据 X_T 的预测。除语义分割模型外,阶段II还在语义分割模型 f_T 的输出空间设置一个判别器 D_{out} ,用于判别输出空间的分割结果属于哪个领域,从而进行对抗学习,增强语义分割模型对于目标域与源域不变特征的提取能力。语义分割模型 f_T 与输出空间判别器 D_{out} 的组合可以看做一个典型的生成对抗网络模型:其中 f_T 视为传统生成对抗网络中的生成器,对 $X_{S \rightarrow T}$ 和 X_T 生成预测结果,同时模型的输出空间设置一个语义判别器 D_{out} 来判断 f_T

输出的结果是来自 $X_{S \rightarrow T}$ 还是 X_T 。

[0051] (1) 子模型网络结构设置

[0052] 整个生成式残差-重采样对抗学习领域自适应框架可以适用于任何语义分割模型,因此可以根据实际任务对于语义分割网络 f_T 进行选择。在方法中, f_T 使用Deeplabv3+语义分割模型作为分割模型基准。为了加速收敛速度,主干特征提取网络 (backbone) 采用深度相对较浅的ResNet-34,并采用了ImageNet预训练权重。输出空间判别器 D_{out} 则为同样为全卷积网络,卷积核大小为 4×4 ,步长为2,卷积核深度为64、128、256、512、1。除最后一层外,各层均包含负斜率为0.2的Leaky ReLU的激活函数。最后一层上采样为输入数据的大小。通过这种方式,语义分割模型能够学习到目标域数据的和源域数据的领域不变的分布特征,从而提升模型在目标域数据的性能表现。

[0053] (2) 子模型目标函数优化

[0054] 判别器 D_{out} 的训练使用了一个二值交叉熵损失 L_{out} :

$$[0055] \quad L_{out} = -\mathbb{E}_{x_T \sim X_T} (\log(1 - D_{out}(f_T(x_T)))) - \mathbb{E}_{x_{S \rightarrow T} \sim X_{S \rightarrow T}} (\log(D_{out}(f_T(x_{S \rightarrow T}))))$$

[0056] 由于 $X_{S \rightarrow T}$ 的尺度为 $H_T \times W_T \times B$,而标记数据 Y_S 的尺度为 $H_S \times W_S \times C$ 。因此使用最近邻插值法用于统一 Y_S 的尺度为 $Y_{S \rightarrow T} \in \mathbb{Z}^{H_T \times W_T \times B}$ 。使用最近邻差值法用于尺度缩放的原因在于其在运算过程中将最邻近的像素赋给原像素点,因此所有的插值均来自原图像,从而不会脱离标记数据中类别标签值的范围。语义分割模型 f_T 的训练使用一个交叉熵损失:

$$[0057] \quad L_{seg} = -\mathbb{E}_{(x_{S \rightarrow T}, y_{S \rightarrow T}) \sim (X_{S \rightarrow T}, Y_{S \rightarrow T})} \left(\sum_{c=1}^C y_{S \rightarrow T} \log(f_T(x_{S \rightarrow T})) \right)$$

[0058] 其中 $y_{S \rightarrow T} \in Y_{S \rightarrow T}$ 为 y_S 经过最近邻插值后统一为目标域尺度后的标记数据,作为 $x_{S \rightarrow T}$ 的标签对模型监督训练。目标域数据 X_T 数据经过 f_T 获得输出结果 $f_T(X_T)$,再输入到判别器 D_{out} 进行对抗学习,对抗损失为:

$$[0059] \quad L_{adv_o} = \mathbb{E}_{x_T \sim X_T} (\log(1 - D_{out}(f_T(x_T))))$$

[0060] 综上,阶段II语义分割模型训练的总损失函数为:

$$[0061] \quad L_{total} = \lambda_{seg} L_{seg} + \lambda_{adv_o} L_{adv_o}$$

[0062] 其中 λ_{seg} 和 λ_{adv_o} 分别为语义分割模型交叉熵损失函数和输出空间对抗学习判别损失函数的超参数。阶段II训练后,最终得到一个跨领域的无人机影像语义分割模型。

[0063] 本领域内的技术人员应明白,本发明的实施例可提供为方法、系统、或计算机程序

产品。因此,本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机、可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0064] 本发明是参照根据本发明实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0065] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0066] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0067] 将上述方法步骤汇编成程序再存储于硬盘或其他非暂态记录媒体,并与处理器电连接即构成了本发明的“一种可读记录媒体”技术方案;而将该可读记录媒体与数据采集装置电连接,通过数据处理能完成遥感图像深度学习模型的构建,则构成本发明的“一种遥感图像深度学习模型的构建系统”技术方案。

[0068] 最后应说明的是:以上所述仅为本发明的优选实施例而已,并不用于限制本发明,尽管参照前述实施例对本发明进行了详细的说明,对于本领域的技术人员来说,其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

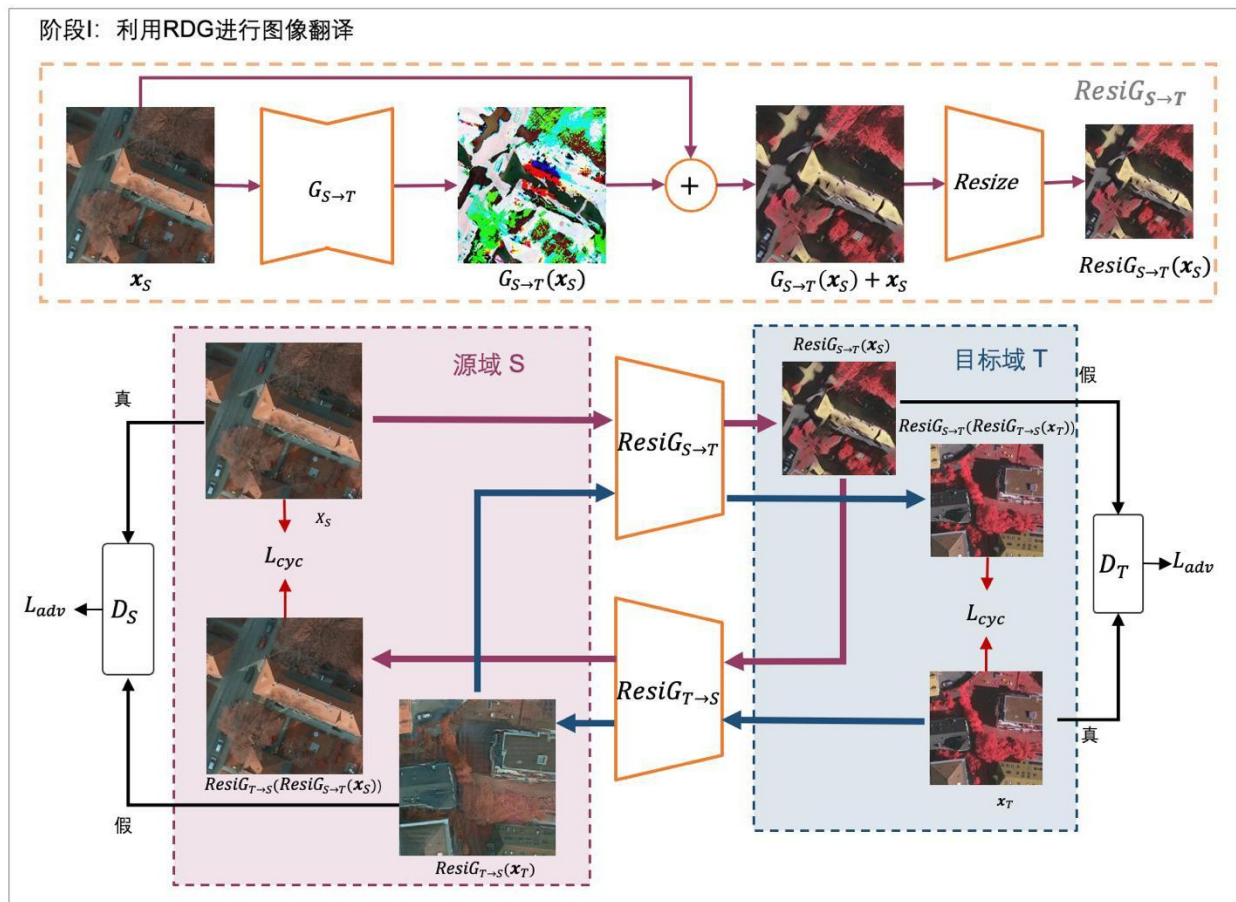


图 1

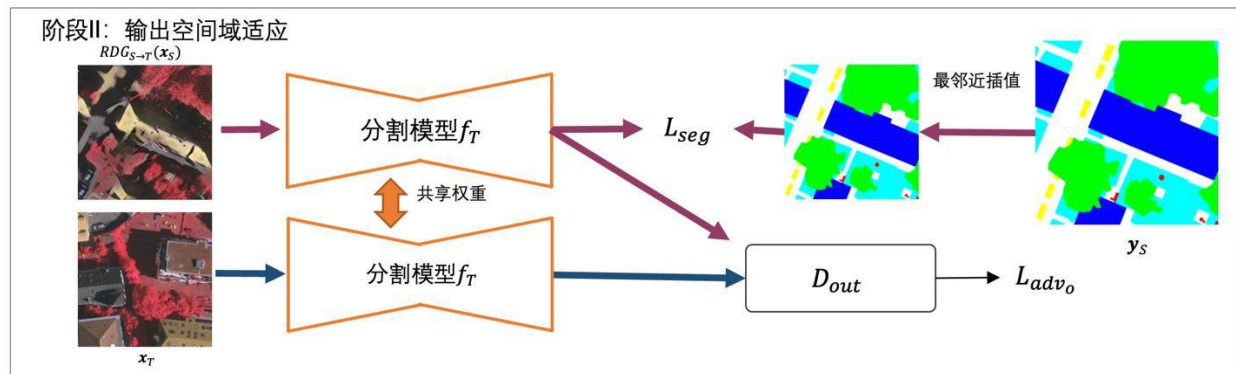


图 2