



(12) 发明专利申请

(10) 申请公布号 CN 116434077 A

(43) 申请公布日 2023. 07. 14

(21) 申请号 202310283534.9

G06V 10/82 (2022.01)

(22) 申请日 2023.03.22

G06N 3/0464 (2023.01)

(71) 申请人 南京信息工程大学

G06N 3/048 (2023.01)

地址 210044 江苏省南京市江北新区宁六路219号

G06N 3/08 (2023.01)

(72) 发明人 夏旻 戴鑫

(74) 专利代理机构 南京苏高专利商标事务所 (普通合伙) 32204

专利代理师 刘红阳

(51) Int. Cl.

G06V 20/13 (2022.01)

G06V 10/26 (2022.01)

G06V 10/44 (2022.01)

G06V 10/80 (2022.01)

G06V 10/764 (2022.01)

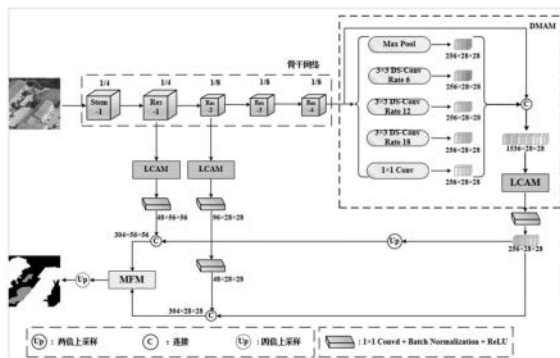
权利要求书3页 说明书8页 附图6页

(54) 发明名称

多尺度位置注意聚合的土地覆盖检测方法、系统和设备

(57) 摘要

本发明公开了一种多尺度位置注意聚合的土地覆盖检测方法、系统和设备。涉及图像处理技术领域,获取训练数据,然后构建多尺度位置注意聚合网络;再训练多尺度位置注意聚合网络,对训练数据中的图像进行数据增强处理,然后将图像与对应的标签转化为张量,然后输入到多尺度位置注意聚合网络中训练;最后使用训练好的多尺度位置注意聚合网络进行预测,经过训练好的权重提取特征进行编码解码操作后输出得到建筑物和水域的掩膜图像;能够有效的减少复杂背景的干扰造成的误分类现象,增强了对建筑物和水域的边缘检测能力,能够精确检测出建筑物和水域区域,避免建筑物和水域相似的光谱特性对检测结果的影响,并且在分割其他物体方面也具有不错的效果。



1. 一种多尺度位置注意聚合的土地覆盖检测方法,其特征在于,所述检测方法包括以下步骤:

获取训练数据,所述训练数据中包括背景、建筑物和水域;

构建多尺度位置注意聚合网络,所述多尺度位置注意聚合网络包括:位置通道注意模块,双支路多尺度模块,多尺度聚合模块;

训练多尺度位置注意聚合网络,对训练数据中的图像进行数据增强处理,然后将图像与对应的标签转化为张量,然后输入到多尺度位置注意聚合网络中训练;

使用训练好的多尺度位置注意聚合网络进行预测,输入一张彩色图片,经过训练好的权重提取特征进行编码解码操作后输出得到建筑物和水域的掩膜图像。

2. 根据权利要求1所述的一种多尺度位置注意聚合的土地覆盖检测方法,其特征在于,所述位置通道注意模块用于将位置信息嵌入通道注意与位置空间注意中,在位置注意中,分别沿着两个空间方向(H和W)进行聚合特征;所述位置通道注意模块的计算公式如下:

$$F_{out} = (M_S + (M_C + x_i))$$

其中 $F_{out} \in \mathbb{R}^{C \times H \times W}$ 表示最终输出, $M_S(\cdot)$ 表示位置注意力操作, $M_C(\cdot)$ 表示通道注意力操作;

通道注意力模块使用了全局平均池化和全局最大值池化两种池化层来提取高层特征,通道注意力模块的计算公式如下:

$$y_{out} = x_i \cdot \text{Sigmoid}(\text{Conv}(\text{Cat}(x_{max}, x_{avg})))$$

其中Sigmoid表示激活函数,Cat表示对两个全局池化输出进行拼接, x_i 为该模块的输入特征信息, x_{max} 和 x_{avg} 分别代表全局最大值池化操作和全局平均池化操作,其计算公式如下:

$$x_{max} = \text{Conv}(g_{max}(x_i))$$

$$x_{avg} = \text{Conv}(g_{avg}(x_i))$$

位置空间注意力模块分别沿着两个空间方向聚合特征,即H和W方向,计算过程如下:

$$x_{i+1} = \text{Conv}\{\sigma(\text{Conv}[\text{Cat}(x_{avg}(x_i), y_{avg}(x_i)^T)])\}$$

x_i 表示输入特征, x_{avg} 表示对 x_i 沿着H方向做自适应平均池化,压缩W的尺寸, y_{avg} 表示对 x_i 沿着W方向做自适应平均池化,压缩H的尺寸, y_{avg}^T 表示将 y_{avg} 进行转置,方便之后的融合操作,Cat表示将 x_{avg} 和 y_{avg}^T 进行融合,汇聚特征信息, $\sigma(\cdot)$ 表示进行BatchNorm和激活函数为ReLU;

汇总H方向和W方向的位置信息,同时获取两个方向的长远距离依赖关系,也保留了两个方向的精确位置信息,之后和原来的输入特征图相结合,构成完整的残差位置空间注意模块;计算过程如下:

$$H_{out} = \sigma(\text{Conv}_g^1(h_i) + \{\text{Softmax}[(\text{Conv}_\Phi^2(h_i) \otimes \text{Conv}_\Theta^{2T}(h_i)) + \text{Conv}_g^1(h_i)]\})$$

$$W_{out} = \sigma(\text{Conv}_g^1(w_i) + \{\text{Softmax}[(\text{Conv}_\Phi^2(w_i) \otimes \text{Conv}_\Theta^{2T}(w_i)) + \text{Conv}_g^1(w_i)]\})$$

$$X_{out} = H_{out} \otimes W_{out}$$

$$X_{location} = X_{out} + X_i$$

其中 h_i 表示H方向的输入注意特征图, w_i 表示W方向的输入注意特征图, Conv^1 和 Conv^2 表示卷积核为1的卷积运算,区别在于通道数不同,Softmax表示激活函数Softmax, $\sigma(\cdot)$ 表示归一化函数BatchNorm, \otimes 表示点乘, X_{out} 表示H方向和W方向的汇总信息, $X_{location}$ 表示位置

空间注意模块的输出, x_i 表示该模块的输入特征图。

3. 根据权利要求1所述的一种多尺度位置注意聚合的土地覆盖检测方法, 其特征在于, 所述双支路多尺度模块的计算方式如下:

$$\begin{aligned} x_{i1} &= \text{Max}(x_i) \\ x_{i2} &= \text{Ds-Conv}_{r=6}(x_i) \\ x_{i3} &= \text{Ds-Conv}_{r=12}(x_i) \\ x_{i4} &= \text{Ds-Conv}_{r=18}(x_i) \\ x_{i5} &= \text{Conv}_{1 \times 1}(x_i) \end{aligned}$$

其中Max表示自适应最大值池化, DS-Conv_r = 分别表示空洞率r=6, 12, 18的空洞卷积, Conv_{1×1}表示卷积核为1的卷积操作。

4. 根据权利要求1所述的一种多尺度位置注意聚合的土地覆盖检测方法, 其特征在于, 所述多尺度融合模块通过以下步骤获得:

首先, 对低层特征和高层特征使用3×3的深度可分离卷积来进一步挖掘特征信息得到 U'_1 和 U'_2 , 加强了语义信息特征;

对于高层特征, 利用双线性插值方法将它采样到与低层特征图相同的大小 U'_{21} , 然后对低层特征进行加权得到 H_1 , 生成对低层特征图增强语义信息的指导;

同时, 低层特征图经过平均池化下采样为高层特征图的尺寸 U'_{12} , 然后加权到高层特征图上, 达到对高层特征图空间信息增强的目的, 接着, 将受到指导的高层特征和低层特征在通道维度上进行融合, 通过增加通道数来实现较好的性能, 这样有利于增加描述图像本身的特征, 计算过程如下:

$$\begin{aligned} U'_1 &= \text{DS-Conv}(U_1) \\ U'_2 &= \text{DS-Conv}(U_2) \\ U'_{12} &= \text{Max}_{3 \times 3}(U'_1) \\ U'_{21} &= \text{Upsample}(U'_2) \\ H_1 &= U'_{11} \otimes \text{Sigmoid}(U'_{21}) \\ L_1 &= \text{Upsample}[U'_{12} \otimes \text{Sigmoid}(U'_{22})] \end{aligned}$$

其中, DS-Conv表示3×3的深度可分离卷积运算, Max_{3×3}是3×3的最大值池化操作, Upsample是两倍上采样操作, \otimes 表示点乘, U_1 为低层特征, U_2 为高层特征, L_1 为 U'_{22} 经过两倍上采样处理的高层特征;

最后, 将融合的特征输入进细化模块, 该模块可以直接嵌入在多尺度融合模块中, 实现端到端的训练, 直接生成精细的分割图, 同时减少了无关区域的干扰; 特征图通过细化模块后再经过四倍的双线性插值上采样恢复成原图尺寸。

5. 根据权利要求1所述的一种多尺度位置注意聚合的土地覆盖检测方法, 其特征在于, 所述背景包括乡村、别墅、工厂、湖泊。

6. 根据权利要求1所述的一种多尺度位置注意聚合的土地覆盖检测方法, 其特征在于, 所述建筑物包括别墅、工厂、平房、高层楼房。

7. 根据权利要求1所述的一种多尺度位置注意聚合的土地覆盖检测方法, 其特征在于, 所述水域由湖泊、泳池和海洋构成。

8. 一种多尺度位置注意聚合的土地覆盖检测系统,其特征在於,所述系统包括:

训练数据模块,所述训练数据中包括背景、建筑物和水域;

多尺度位置注意聚合网络模块,所述多尺度位置注意聚合网络模块包括:位置通道注意模块,双支路多尺度模块,多尺度聚合模块;

训练多尺度位置注意聚合网络模块,所述训练多尺度位置注意聚合网络模块用于对训练数据中的图像进行数据增强处理,然后将图像与对应的标签转化为张量,然后输入到多尺度位置注意聚合网络中训练;

预测模块:所述预测模块通过训练好的多尺度位置注意聚合网络进行预测,输入一张彩色图片,经过训练好的权重提取特征进行编码解码操作后输出得到建筑物和水域的掩膜图像。

9. 一种设备,其特征在於,包括:

一个或多个处理器;

存储器,用于存储一个或多个程序;

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-7中任一所述的一种多尺度位置注意聚合的土地覆盖检测方法。

10. 一种包含计算机可执行指令的存储介质,其特征在於,所述计算机可执行指令在由计算机处理器执行时用于执行如权利要求1-7中任一所述的一种多尺度位置注意聚合的土地覆盖检测方法。

多尺度位置注意聚合的土地覆盖检测方法、系统和设备

技术领域

[0001] 本发明涉及图像处理技术领域,尤其涉及一种多尺度位置注意聚合的土地覆盖检测方法、系统和设备。

背景技术

[0002] 通过分析土地覆盖数据集,可以了解城市化率、水覆盖率和其它环境变化。高分辨率遥感卫星技术的快速发展对地表分类的精度提出了挑战,近年来人工智能等新技术的发展为图像自动分割提供了实现途径。在遥感图像中,水域和建筑物是非常重要的地标,它们在土地规划、土地分类、城镇化速度研究、水资源管理和水资源保护等方面具有重要的现实意义,因此需要对水域与建筑物进行精确分割;土地覆盖检测方法可分为阈值方法和深度学习方法。阈值方法大多是基于建筑物与水域的光谱特性,特征分析能力有限,面对高分辨率图像,分割效果较差。而深度学习因为下采样过深的原因,导致特征图的信息损失严重,图像位置信息发生了偏移,割裂了整体效果,直接运用于土地覆盖任务,容易丢失空间信息,造成网络误分类情况;为此,现在提出一种多尺度位置注意模块的土地覆盖检测方法、设备及存储介质。

发明内容

[0003] 发明目的:本发明的目的是提供一种分割效果好、分类结果准确的多尺度位置注意聚合的土地覆盖检测方法;本发明的另一目的是提供一种多尺度位置注意聚合的土地覆盖检测系统。

[0004] 技术方案:本发明的一种多尺度位置注意聚合的土地覆盖检测方法,所述检测方法包括以下步骤:

[0005] 获取训练数据,所述训练数据中包括背景、建筑物和水域;

[0006] 构建多尺度位置注意聚合网络,所述多尺度位置注意聚合网络包括:位置通道注意模块,双支路多尺度模块,多尺度聚合模块;

[0007] 训练多尺度位置注意聚合网络,对训练数据中的图像进行数据增强处理,然后将图像与对应的标签转化为张量,然后输入到多尺度位置注意聚合网络中训练;

[0008] 使用训练好的多尺度位置注意聚合网络进行预测,输入一张彩色图片,经过训练好的权重提取特征进行编码解码操作后输出得到建筑物和水域的掩膜图像。

[0009] 作为上述方案的进一步改进,所述位置通道注意模块用于将位置信息嵌入通道注意与位置空间注意中,在位置注意中,分别沿着两个空间方向(H和W)进行聚合特征;所述位置通道注意模块的计算公式如下:

$$F_{out} = (M_S + (M_C + x_i))$$

[0011] 其中 $F_{out} \in R^{C \times H \times W}$ 表示最终输出, $M_S()$ 表示位置注意力操作, $M_C()$ 表示通道注意力操作;

[0012] 通道注意力模块使用了全局平均池化和全局最大值池化两种池化层来提取高层

特征,通道注意力模块的计算公式如下:

$$[0013] \quad y_{\text{out}} = x_i \cdot \text{Sigmoid}(\text{Conv}(\text{Cat}(x_{\text{max}}, x_{\text{avg}})))$$

[0014] 其中Sigmoid表示激活函数,Cat表示对两个全局池化输出进行拼接, x_i 为该模块的输入特征信息, x_{max} 和 x_{avg} 分别代表全局最大值池化操作和全局平均池化操作,其计算公式如下:

$$[0015] \quad x_{\text{max}} = \text{Conv}(g_{\text{max}}(x_i))$$

$$[0016] \quad x_{\text{avg}} = \text{Conv}(g_{\text{avg}}(x_i))$$

[0017] 位置空间注意力模块分别沿着两个空间方向聚合特征,即H和W方向,计算过程如下:

$$[0018] \quad x_{i+1} = \text{Conv}\{\sigma(\text{Conv}[\text{Cat}(x_{\text{avg}}(x_i), y_{\text{avg}}(x_i)^T])\}$$

[0019] x_i 表示输入特征, x_{avg} 表示对 x_i 沿着H方向做自适应平均池化,压缩W的尺寸, y_{avg} 表示对 x_i 沿着W方向做自适应平均池化,压缩H的尺寸, y_{avg}^T 表示将 y_{avg} 进行转置,方便之后的融合操作,Cat表示将 x_{avg} 和 y_{avg} 进行融合,汇聚特征信息, $\sigma(\cdot)$ 表示进行BatchNorm和激活函数为ReLU;

[0020] 汇总H方向和W方向的位置信息,同时获取两个方向的长远距离依赖关系,也保留了两个方向的精确位置信息,之后和原来的输入特征图相结合,构成完整的残差位置空间注意模块;计算过程如下:

$$[0021] \quad H_{\text{out}} = \sigma(\text{Conv}_g^1(h_i) + \{\text{Softmax}[(\text{Conv}_\Phi^2(h_i) \otimes \text{Conv}_\Theta^{2T}(h_i)) + \text{Conv}_g^1(h_i)]\})$$

$$[0022] \quad W_{\text{out}} = \sigma(\text{Conv}_g^1(w_i) + \{\text{Softmax}[(\text{Conv}_\Phi^2(w_i) \otimes \text{Conv}_\Theta^{2T}(w_i)) + \text{Conv}_g^1(w_i)]\})$$

$$[0023] \quad X_{\text{out}} = H_{\text{out}} \otimes W_{\text{out}}$$

$$[0024] \quad X_{\text{location}} = X_{\text{out}} + X_i$$

[0025] 其中 h_i 表示H方向的输入注意特征图, w_i 表示W方向的输入注意特征图, Conv^1 和 Conv^2 表示卷积核为1的卷积运算,区别在于通道数不同,Softmax表示激活函数Softmax, $\sigma(\cdot)$ 表示归一化函数BatchNorm, \otimes 表示点乘, X_{out} 表示H方向和W方向的汇总信息, X_{location} 表示位置空间注意模块的输出, X_i 表示该模块的输入特征图。

[0026] 作为上述方案的进一步改进,所述双支路多尺度模块的计算方式如下:

$$[0027] \quad x_{i1} = \text{Max}(x_i)$$

$$[0028] \quad x_{i2} = \text{Ds-Conv}_{r=6}(x_i)$$

$$[0029] \quad x_{i3} = \text{Ds-Conv}_{r=12}(x_i)$$

$$[0030] \quad x_{i4} = \text{Ds-Conv}_{r=18}(x_i)$$

$$[0031] \quad x_{i5} = \text{Conv}_{1 \times 1}(x_i)$$

[0032] 其中Max表示自适应最大值池化,DS-Conv_{r=}分别表示空洞率 $r=6,12,18$ 的空洞卷积, $\text{Conv}_{1 \times 1}$ 表示卷积核为1的卷积操作。

[0033] 作为上述方案的进一步改进,所述多尺度融合模块通过以下步骤获得:

[0034] 首先,对低层特征和高层特征使用 3×3 的深度可分离卷积来进一步挖掘特征信息得到 U'_1 和 U'_2 ,加强了语义信息特征;

[0035] 对于高层特征,利用双线性插值方法将它采样到与低层特征图相同的大小 U'_{21} ,然

后对低层特征进行加权得到 H_1 ,生成对低层特征图增强语义信息的指导;

[0036] 同时,低层特征图经过平均池化下采样为高层特征图的尺寸 U'_{12} ,然后加权到高层特征图上,达到对高层特征地图空间信息增强的目的,接着,将受到指导的高层特征和低层特征在通道维度上进行融合,通过增加通道数来实现较好的性能,这样有利于增加描述图像本身的特征,计算过程如下:

$$[0037] \quad U'_1 = \text{DS-Conv}(U_1)$$

$$[0038] \quad U'_2 = \text{DS-Conv}(U_2)$$

$$[0039] \quad U'_{12} = \text{Max}_{3 \times 3}(U'_1)$$

$$[0040] \quad U'_{21} = \text{Upsample}(U'_2)$$

$$[0041] \quad H_1 = U'_{11} \otimes \text{Sigmoid}(U'_{21})$$

$$[0042] \quad L_1 = \text{Upsample}[U'_{12} \otimes \text{Sigmoid}(U'_{22})]$$

[0043] 其中,DS-Conv表示 3×3 的深度可分离卷积运算, $\text{Max}_{3 \times 3}$ 是 3×3 的最大值池化操作,Upsample是两倍上采样操作, \otimes 表示点乘, U_1 为低层特征, U_2 为高层特征, L_1 为 U'_{22} 经过两倍上采样处理的高层特征;

[0044] 最后,将融合的特征输入进细化模块,该模块可以直接嵌入在多尺度融合模块中,实现端到端的训练,直接生成精细的分割图,同时减少了无关区域的干扰;特征图通过细化模块后再经过四倍的双线性插值上采样恢复成原图尺寸。

[0045] 优选的,所述背景包括乡村、别墅、工厂、湖泊。

[0046] 优选的,所述建筑物包括别墅、工厂、平房、高层楼房。

[0047] 优选的,所述水域由湖泊、泳池和海洋构成。

[0048] 另一方面,本发明提供一种多尺度位置注意聚合的土地覆盖检测系统,所述系统包括:

[0049] 训练数据模块,所述训练数据中包括背景、建筑物和水域;

[0050] 多尺度位置注意聚合网络模块,所述多尺度位置注意聚合网络模块包括:位置通道注意模块,双支路多尺度模块,多尺度聚合模块;

[0051] 训练多尺度位置注意聚合网络模块,所述训练多尺度位置注意聚合网络模块用于对训练数据中的图像进行数据增强处理,然后将图像与对应的标签转化为张量,然后输入到多尺度位置注意聚合网络中训练;

[0052] 预测模块:所述预测模块通过训练好的多尺度位置注意聚合网络进行预测,输入一张彩色图片,经过训练好的权重提取特征进行编码解码操作后输出得到建筑物和水域的掩膜图像。

[0053] 另一方面,本发明提供一种设备,包括:

[0054] 一个或多个处理器;

[0055] 存储器,用于存储一个或多个程序;

[0056] 当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如上述的一种多尺度位置注意模块的土地覆盖检测方法。

[0057] 另一方面,本发明提供一种包含计算机可执行指令的存储介质,所述计算机可执行指令在由计算机处理器执行时用于执行如上述的一种多尺度位置注意模块的土地覆盖

检测方法。

[0058] 有益效果：与现有技术相比，本发明具有如下显著优点：本发明在使用的过程中，通过获取训练数据，所述训练数据中包括的背景、建筑物和水域；然后构建多尺度位置注意聚合网络，所述多尺度位置注意聚合网络包括：位置通道注意模块，双支路多尺度模块，多尺度融合模块；再训练多尺度位置注意聚合网络，对训练数据中的图像进行数据增强处理，然后将图像与对应的标签转化为张量，然后输入到多尺度位置注意聚合网络中训练；最后使用训练好的多尺度位置注意聚合网络进行预测，输入一张彩色图片，经过训练好的权重提取特征进行编码解码操作后输出得到建筑物和水域的掩膜图像；能够有效的减少复杂背景的干扰造成的误分类现象，增强了对建筑物和水域的边缘检测能力，能够精确检测出建筑物和水域区域，避免建筑物和水域相似的光谱特性对检测结果的影响，并且在分割其他物体方面也具有不错的效果，泛化能力优秀。

附图说明

[0059] 图1为本发明用于土地覆盖识别的多尺度位置注意聚合网络结构图；

[0060] 图2为本发明的位置通道注意模块示意图；

[0061] 图3为本发明的双支路多尺度模块示意图；

[0062] 图4为本发明的多尺度融合模块示意图；

[0063] 图5为本发明实施例的训练集样本图，其中上半部分为遥感图像原图，下半部分为对应的标签图；

[0064] 图6为本发明实施例进行预测的遥感图像原图；

[0065] 图7为本发明实施例进行预测的遥感图像原图的预测图。

具体实施方式

[0066] 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清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其它实施例，都属于本发明保护的范围。

[0067] 如图1所示，一种多尺度位置注意聚合的土地覆盖辨识方法，包括：

[0068] 训练数据集的获取：

[0069] 数据集来自于Google Earth的分辨率为 1600×900 的地图，它包含了欧洲的乡村公园，中国的沿海居民区，北美的私人别墅等，具有空间跨度大、拍摄角度多样化、视角低、环境复杂等特点，因此该数据集对算法要求较高。本文将这些大图像分割成大小为 224×224 的图像，并对这些图片进行数据增强处理，获得10000对包含建筑物与水体的图像与标签。数据增强的策略包括三种类型：水平翻转（50%）、垂直翻转（50%）、随机旋转（ -10° 至 10° ）等。训练集和数据集的比例为8:2。这些图片被标记为3类对象：水域、建筑物和背景。该数据集具有以下特点：(1) 如图5(a)和图5(d)所示，图像中的集装箱房屋的顶部以及一些杂物与水域和背景之间在像素值上有相似性，容易使模型产生误分类，对模型的检测能力具有挑战性；(2) 如图5(a)和图5(c)所示，由于拍摄角度不同，图像中的建筑物有一定的位置偏差，而不是单一的视角图像；(3) 土地覆盖的场景较多，能够真实的反映模型的分割能力；

(4) 房屋之间存在高度差,高层建筑阴影投影映射在低层建筑上会干扰模型分割;(5) 我们使用的是人工标注,因此分割区域的信息较为精确。如图5所示,展示了LandCover数据集集中的部分训练图像,图中第一行是原始彩色图像,第二行是对应的标签,其中白色区域代表建筑物,蓝色区域代表水域,黑色区域代表背景。

[0070] 构建多尺度位置注意聚合网络

[0071] 一种基于多尺度位置注意聚合网络的遥感图像土地覆盖的检测方法的结构如图1所示。整个网络是一个编码器-解码器的结构,采用端到端的训练方式,主要由位置通道注意模块,双支路多尺度模块和多尺度聚合模块构成。由于在对遥感图像中识别土地覆盖的过程中,是对图中特征信息的提取十分重要,选择合适的主干网络可以大大提高网络的检测效率,本发明选取的是残差网络作为主干网络来进行特征提取操作。

[0072] 所述的一种多尺度位置注意聚合的土地覆盖检测方法,其特征在于,所述位置注意模块用于将位置信息嵌入通道注意与位置空间注意中,在位置注意中,分别沿着两个空间方向(H和W)进行聚合特征,即不仅专注的捕获特征图一个方向的远程依赖关系,同时也可以保留该方向精确的位置信息,加强了网络对于感兴趣对象的捕捉。位置全局注意模块实现了捕获像素位置之间的长距离依赖性关系,更好的获取全局信息。通道注意模块也有有效的关注了建筑物与水域的边缘信息,获取有效的细节信息,最后将通道注意模块和位置空间注意模块的特征信息进行相加得到最后的输出;整个模块的计算过程如下:

[0073] $F_{out} = (M_S + (M_C + x_i))$

[0074] 其中 $F_{out} \in R^{C \times H \times W}$ 表示最终输出, $M_S()$ 表示位置注意力操作, $M_C()$ 表示通道注意力操作;

[0075] 通道注意力模块使用了全局平均池化和全局最大值池化两种池化层来提取高层特征。不同类型的全局池化层可以提取丰富的高层特征。特征信息经过两个全局池化层后,通过卷积核为1的卷积层进行通道转换,然后将两个支路的输出进行拼接。接下来,使用卷积核为1的卷积作为选择器,自适应地聚焦于两个全局池化的特征表示。它的计算公式如下:

[0076] $y_{out} = x_i \cdot \text{Sigmoid}(\text{Conv}(\text{Cat}(x_{max}, x_{avg})))$

[0077] 其中Sigmoid表示激活函数,Cat表示对两个全局池化输出进行拼接, x_i 为该模块的输入特征信息, x_{max} 和 x_{avg} 分别代表全局最大值池化操作和全局平均池化操作,其计算公式如下:

[0078] $x_{max} = \text{Conv}(g_{max}(x_i))$

[0079] $x_{avg} = \text{Conv}(g_{avg}(x_i))$

[0080] 位置空间注意力模块分别沿着两个空间方向聚合特征,即H和W方向。位置注意将输入特征通过自适应池化层分解为两个1D特征编码过程,以有效地将空间信息整合到生成的注意力特征图中。我们先将输入的特征信息分别沿着H和W方向做自适应平均池化操作,接着将H和W方向的特征信息进行融合,汇聚特征信息,计算过程如下:

[0081] $x_{i+1} = \text{Conv}\{\sigma(\text{Conv}[\text{Cat}(x_{avg}(x_i), y_{avg}(x_i)^T])\})\}$

[0082] x_i 表示输入特征, x_{avg} 表示对 x_i 沿着H方向做自适应平均池化,压缩W的尺寸, y_{avg} 表示对 x_i 沿着W方向做自适应平均池化,压缩H的尺寸, y_{avg}^T 表示将 y_{avg} 进行转置,方便之后的

融合操作,Cat表示将 x_{avg} 和 y_{avg} 进行融合,汇聚特征信息, $\sigma(\cdot)$ 表示进行BatchNorm和激活函数为ReLU。

[0083] 通过上述操作后的特征图将会被拆分,之后被分别送进H和W方向的全局注意力模块中。以H方向的位置注意为例,我们将输入进H方向的特征信息平均分为三个位置元素 (Θ, Φ, g) , Θ 和 Φ 分别进行通道合并及转置操作后,进行矩阵相乘,它们计算的结果经过Softmax后,再和经过通道变化后的 g 进行矩阵相乘。通过三个位置元素之间运算,我们可以实现位置信息的交互,同时通过Softmax操作对经过全局注意后的特征图重新加权,与带有原特征图信息的 g 相加,补全信息,从而捕捉特征图的长远距离依赖,并且保留了H方向的精确信息,W方向的位置注意操作也相同。最后,我们汇总H方向和W方向的位置信息,同时获取两个方向的长远距离依赖关系,也保留了两个方向的精确位置信息,之后和原来的输入特征图相结合,构成完整的残差位置空间注意模块;计算过程如下:

$$[0084] \quad H_{out} = \sigma(\text{Conv}_g^1(h_i) + \{\text{Softmax}[(\text{Conv}_\Phi^2(h_i) \otimes \text{Conv}_\Theta^{2T}(h_i)) + \text{Conv}_g^1(h_i)]\})$$

$$[0085] \quad W_{out} = \sigma(\text{Conv}_g^1(w_i) + \{\text{Softmax}[(\text{Conv}_\Phi^2(w_i) \otimes \text{Conv}_\Theta^{2T}(w_i)) + \text{Conv}_g^1(w_i)]\})$$

$$[0086] \quad X_{out} = H_{out} \otimes W_{out}$$

$$[0087] \quad X_{location} = X_{out} + X_i$$

[0088] 其中 h_i 表示H方向的输入注意特征图, w_i 表示W方向的输入注意特征图, Conv^1 和 Conv^2 表示卷积核为1的卷积运算,区别在于通道数不同,Softmax表示激活函数Softmax, $\sigma(\cdot)$ 表示归一化函数BatchNorm, \otimes 表示点乘, X_{out} 表示H方向和W方向的汇总信息, $X_{location}$ 表示位置空间注意模块的输出, X_i 表示该模块的输入特征图。

[0089] 需要进一步进行说明的是,所述双支路多尺度模块用于提取目标的边缘信息与不同尺度的特征,并通过空洞卷积的方式增大感受野,进一步的提取多尺度上下文信息;并行的空洞卷积层使得每个采样率的特征图都能在单独的分支中得到进一步处理,互不干扰,这有利于最后多尺度信息的融合。另一方面,经过类似残差连接通道的特征图,在经过卷积核为1的卷积运算后,与完成并行空洞卷积的特征图融合,补充卷积运算后可能丢失的原始信息,从而实现精准分割;其计算方式如下:

$$[0090] \quad x_{i1} = \text{Max}(x_i)$$

$$[0091] \quad x_{i2} = \text{Ds-Conv}_{r=6}(x_i)$$

$$[0092] \quad x_{i3} = \text{Ds-Conv}_{r=12}(x_i)$$

$$[0093] \quad x_{i4} = \text{Ds-Conv}_{r=18}(x_i)$$

$$[0094] \quad x_{i5} = \text{Conv}_{1 \times 1}(x_i)$$

[0095] 其中Max表示自适应最大值池化,DS-Conv_{r=}分别表示空洞率 $r=6,12,18$ 的空洞卷积,Conv_{1×1}表示卷积核为1的卷积操作。

[0096] 需要进一步进行说明的是,在具体实施过程中,所述的一种多尺度位置注意模块的土地覆盖检测方法,其特征在于,所述多尺度融合模块:本发明设计了多尺度融合模块用于高层语义信息和低层空间信息进行相互指导融合,满足土地覆盖分割任务中复杂的细小支流和房屋房檐边缘的分割要求,通过不同尺度的指导,增强通道之间的信息交互能力。首先对低层特征和高层特征使用 3×3 的深度可分离卷积来进一步挖掘特征信息得到 U'_1 和 U'_2 ,加强了语义信息特征。对于高层特征,利用双线性插值方法将它采样到与低层特征图相

同的大小 U'_{21} ,然后对低层特征进行加权得到 H_1 ,生成对低层特征图增强语义信息的指导。同时,低层特征图经过平均池化下采样为高层特征图的尺寸 U'_{12} ,然后加权到高层特征图上,达到对高层特征图空间信息增强的目的,接着,将受到指导的高层特征和低层特征在通道维度上进行融合,通过增加通道数来实现较好的性能,这样有利于增加描述图像本身的特征,计算过程如下:

$$[0097] \quad U'_1 = \text{DS-Conv}(U_1)$$

$$[0098] \quad U'_2 = \text{DS-Conv}(U_2)$$

$$[0099] \quad U'_{12} = \text{Max}_{3 \times 3}(U'_1)$$

$$[0100] \quad U'_{21} = \text{Upsample}(U'_2)$$

$$[0101] \quad H_1 = U'_{11} \otimes \text{Sigmoid}(U'_{21})$$

$$[0102] \quad L_1 = \text{Upsample}[U'_{12} \otimes \text{Sigmoid}(U'_{22})]$$

[0103] 其中,DS-Conv表示 3×3 的深度可分离卷积运算, $\text{Max}_{3 \times 3}$ 是 3×3 的最大值池化操作,Upsample是两倍上采样操作, \otimes 表示点乘, U_1 为低层特征, U_2 为高层特征, L_1 为 U'_{22} 经过两倍上采样处理的高层特征。最后,将融合的特征输入进细化模块,该模块可以直接嵌入在多尺度融合模块中,实现端到端的训练,直接生成精细的分割图,同时减少了无关区域的干扰。特征图通过细化模块后再经过四倍的双线性插值上采样恢复成原图尺寸。

[0104] 由于建筑物与水域的类型丰富,存在一些楼体中间镂空的建筑,水池中有观测塔,这些会造成网络的误分类。除此之外,一些建筑物的屋顶,比如蓝顶集装箱,它们在可见光谱中与水域具有相似的属性,这会干扰网络判断,造成误分类,这是由于缺失位置信息和全局信息导致的。在编码阶段,网络获取了足够的语义信息,但是底层的特征图因为多次的下采样以及卷积操作,损失了很多全局信息,导致网络丢失了水域或者建筑物的大尺度特征,因此会将蓝色屋顶的房屋和水域错误分类。为了获得更准确的分类效果,本文设计了一个位置通道注意模块(LCAM),该模块可以将位置信息与全局信息结合。在位置信息中,LCAM增强了网络对房屋镂空边缘,房屋顶部游泳池边缘的关注。全局信息中,LCAM集中获取大型工厂、蓝色屋顶的集装箱房屋的整体特征,以此达到优秀的分割效果。

[0105] 主干网络进行特征提取时,在网络的深层会产生大量的语义特征信息以及类别信息,但是却不是所有的信息都是有用的,如何提取深层语义信息并利用类别信息指导模型对最终的分割结果会产生很大的影响。在土地覆盖检测任务中,传统的提取语义信息的方法对大尺度的会场、湖泊非常有效,但对于小的水域却不是很有效,大的卷积核会提取很多无关的信息,干扰模型的预测,除此之外,该模块的输出通道数过大,导致计算参数量过高,效率较低,基于以上问题,这里提出了一种双支路多尺度模块(DMAM)用于获取水域与建筑物的多尺度信息,并行的空洞卷积层使得每个采样率的特征图都能在单独的分支中得到进一步处理,互不干扰,有利于最后多尺度信息的融合。另一方面,经过类似残差连接通道的特征图,在经过卷积核为1的卷积运算后,与完成并行空洞卷积的特征图融合,补充卷积运算后可能丢失的原始信息,从而实现精准分割。

[0106] 目前大多数的特征融合方法只是简单的维度拼接或者利用广播机制进行相加,这种粗糙的处理方式会导致低层空间信息和高层语义信息多样性的丢失。为了满足土地覆盖分割任务中复杂的细小支流和房屋房檐边缘的分割要求,需要多种特征信息进行融合,因

此多尺度融合模块是必不可少的。针对以上的问题,我们提出了多尺度融合模块,高层特征具有丰富的类别信息,可以指导低层特征进行分类,而低层特征保留着大量的空间信息,在MFM中我们使用多尺度融合模块来完成不同尺度的相互指导,增强通道之间的信息交互能力,可以弥补高层对建筑物与水域整体把握的不足。

[0107] 随着编码器部分输出特征图逐步上采样到原图大小的过程中,我们使用MFM中的细化卷积模块来实现端到端的训练,直接生成精细的分割图,同时减少了无关区域的干扰。特征图通过细化模块后再经过四倍的双线性插值上采样恢复成原图尺寸。

[0108] 在解码过程中如果单一地对特征图进行上采样操作会损失许多特征,造成最终的分割结果的失败,但是简单的线性操作,例如求和或者串联却又不是最佳的选择,无法关注到融合信息中有用的部分。土地覆盖检测任务中各种杂物和建筑物、水域相似的浅层特征以及复杂背景的干扰都要求我们的网络在上采样恢复过程中能够准确提取不同粒度地特征,以融合更丰富的特征信息,从而最小化相似对象的影响。

[0109] 在我们的网络中存在大量的通道连接操作,这样不仅可以融合不同层次的信息,还能提高模型的检测效率,避免信息的遗漏。

[0110] 利用数据集进行网络模型的训练:

[0111] 本发明采用有监督训练的方式,首先对数据集中的图像进行数据增强处理,然后再将原始图片和对应的标签转化为张量后输入到模型当中进行训练。采用的优化器为自适应矩估计(Adam),学习率策略采用的是动态调整学习率“Poly”策略,其中将初始学习率设置为0.00015,衰减系数为0.98,每训练3次更新一次学习率,一共训练了300次,训练时批量大小为16。

[0112] 使用训练好的网络模型进行预测并生成土地覆盖的掩膜图像:

[0113] 在训练结束后会得到模型的权重,接下来进入模型的预测阶段,本发明在预测时,使用训练好的多尺度位置聚合网络进行预测,输入一张原始彩色图片,如图6所示为进行预测的原始彩色图片,这里我们选择的是一张大场景图片,经过训练好的权重提取特征进行编码解码操作后输出便得到土地覆盖的掩膜图像,如图7所示。

[0114] 在本说明书的描述中,参考术语“一个实施例”、“示例”、“具体示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不一定指的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任何的一个或多个实施例或示例中以合适的方式结合。

[0115] 以上显示和描述了本发明的基本原理、主要特征和本发明的优点。本行业的技术人员应该了解,本发明不受上述实施例的限制,上述实施例和说明书中描述的只是说明本发明的原理,在不脱离本发明精神和范围的前提下,本发明还会有各种变化和改进,这些变化和进步都落入要求保护的本发明范围内容。

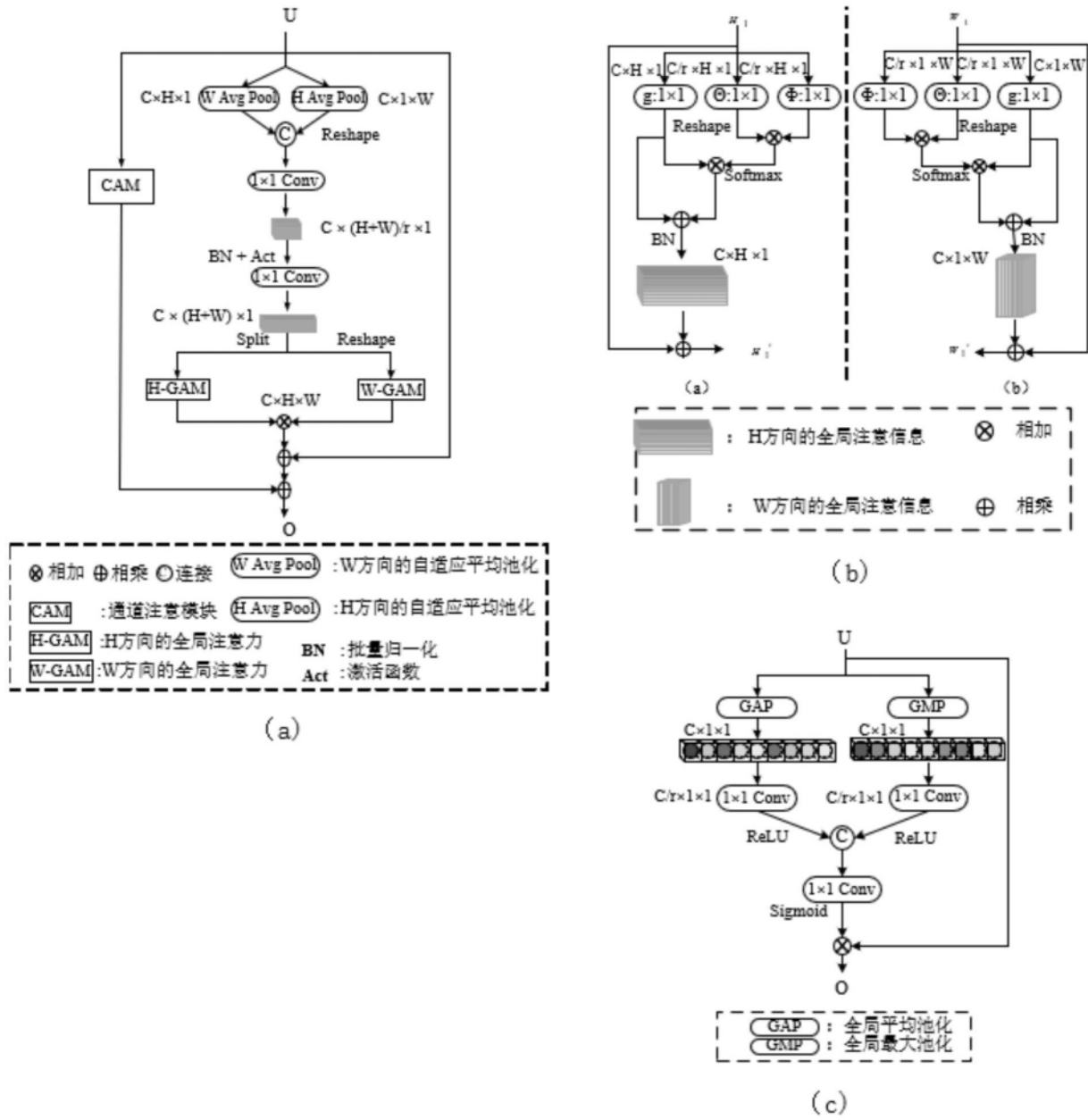


图2

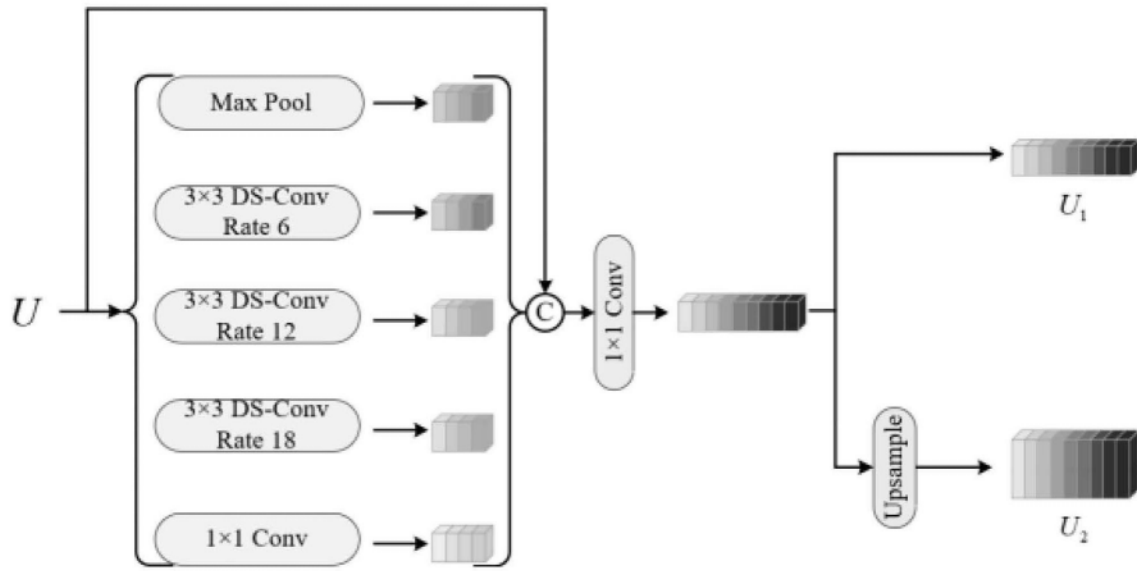


图3

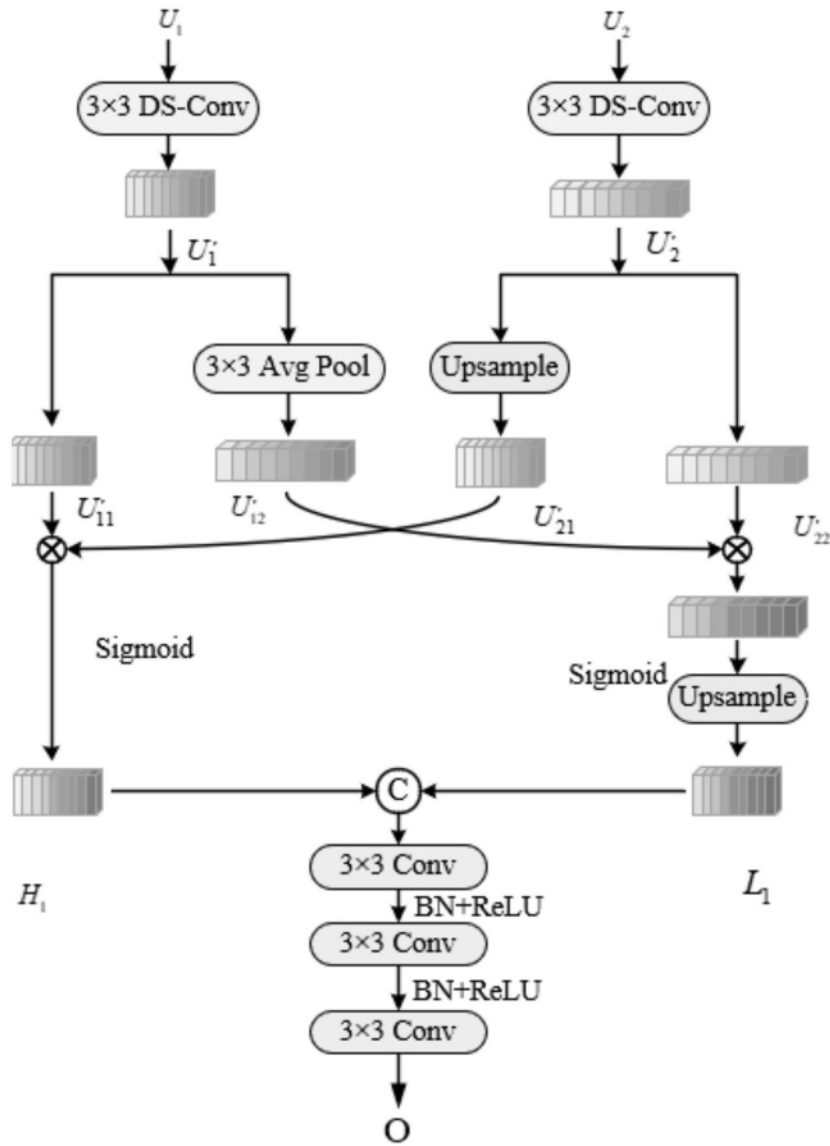


图4

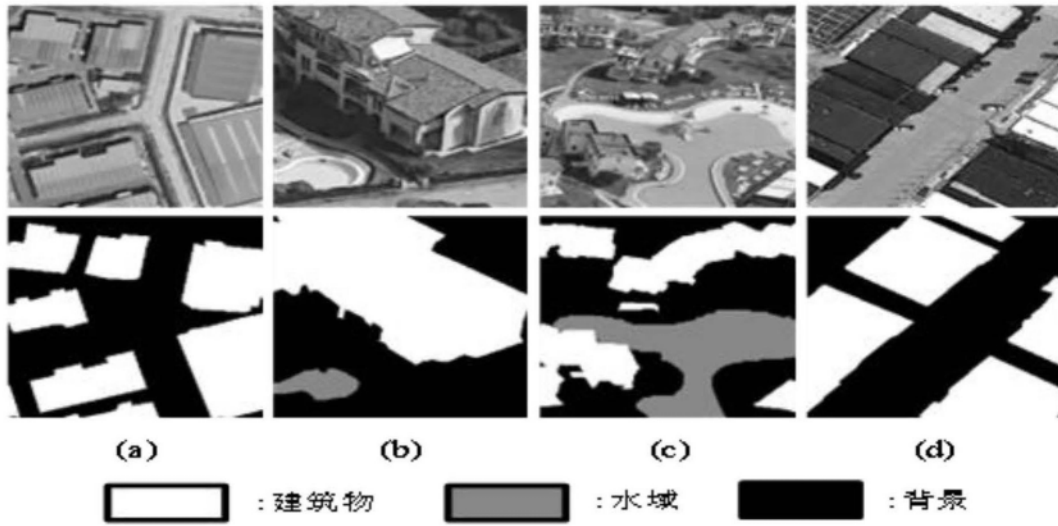


图5



图6



图7