



(12) 发明专利申请

(10) 申请公布号 CN 113435286 A

(43) 申请公布日 2021. 09. 24

(21) 申请号 202110683465.1

G06N 3/08 (2006.01)

(22) 申请日 2021.06.21

(71) 申请人 电子科技大学

地址 611731 四川省成都市高新区(西区)  
西源大道2006号

(72) 发明人 陈怀新 刘壁源 黄周

(74) 专利代理机构 北京正华智诚专利代理事务  
所(普通合伙) 11870

代理人 何凡

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/20 (2006.01)

G06K 9/32 (2006.01)

G06K 9/62 (2006.01)

G06N 3/04 (2006.01)

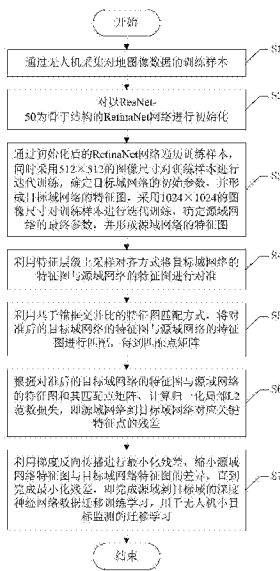
权利要求书2页 说明书4页 附图2页

(54) 发明名称

用于无人机小目标监测的迁移学习方法

(57) 摘要

本发明公开了一种用于无人机小目标监测的迁移学习方法,其包括将无人机采集的开源对地图像数据作为训练样本;通过初始化后的RetinaNet网络遍历训练样本,同时采用不同图像尺寸对训练样本进行训练,得到对应的特征图;利用特征层级上采样对齐方式将对应的特征图进行对准;利用基于锚框交并比的自适应特征图匹配方式,得到对准后的特征图的匹配点矩阵;计算对准后的特征图关键特征点的残差;利用梯度反向传播进行最小化残差,直到完成最小化残差,即完成源域到目标域的深度学习数据迁移训练学习,用于无人机小目标监测的迁移学习。本发明模型收敛效果好、提高了小目标检测的精度、降低了数据制备成本以及避免了计算量的过多增长。



1. 一种用于无人机小目标监测的迁移学习方法, 其特征在于, 包括以下步骤:

S1、将无人机采集的开源对地图像数据作为训练样本;

S2、对以ResNet-50为骨干结构的RetinaNet网络进行初始化;

S3、通过初始化后的RetinaNet网络遍历训练样本, 同时采用 $512 \times 512$ 的图像尺寸对训练样本进行迭代训练, 确定目标域网络的初始参数, 并形成目标域网络的特征图, 采用 $1024 \times 1024$ 的图像尺寸对训练样本进行迭代训练, 确定源域网络的最终参数, 并形成源域网络的特征图;

S4、利用特征层级上采样对齐方式将目标域网络的特征图与源域网络的特征图进行对准;

S5、利用基于锚框交并比的自适应特征图匹配方式, 将对准后的目标域网络与源域网络的层级特征图进行匹配, 得到匹配点矩阵;

S6、根据对准后的目标域网络的特征图与源域网络的特征图和其匹配点矩阵, 计算归一化局部L2范数损失, 即源域网络到目标域网络对应关键特征点的残差;

S7、利用梯度反向传播进行最小化残差, 缩小源域网络特征图与目标域网络特征图的差异, 直到完成最小化残差, 即完成源域到目标域的深度神经网络数据迁移训练学习, 用于无人机小目标监测的迁移学习。

2. 根据权利要求1所述的用于无人机小目标监测的迁移学习方法, 其特征在于: 步骤S1中, 采集的开源对地图像数据为VisDrone2018-DET数据集, 训练样本为6471张。

3. 根据权利要求1所述的用于无人机小目标监测的迁移学习方法, 其特征在于, 步骤S2的具体方法包括以下子步骤:

S2-1、利用超过33万张图片的COCO数据对RetinaNet网络中的ResNet-50结构进行预训练; 其中ResNet-50结构为卷积特征抽取网络;

S2-2、通过随机梯度下降方式对RetinaNet网络进行训练参数设置; 其中初始学习率、动量和权重衰减系数分别设置为0.001、0.9、0.01。

4. 根据权利要求1所述的用于无人机小目标监测的迁移学习方法, 其特征在于: 步骤S3中遍历的次数为20轮, 每次迭代的输入为4张图片压缩的张量, 即每轮的迭代次数为训练样本的张数除以4; 整个20轮遍历中, 前1000次迭代学习率从0.001到0.01线性变化, 除第6轮和第14轮遍历训练样本的学习率为前一轮的 $1/5$ , 后续迭代的学习率与前一次相同。

5. 根据权利要求1所述的用于无人机小目标监测的迁移学习方法, 其特征在于: 步骤S4中特征层级上采样对齐方式为最近邻插值方式。

6. 根据权利要求1所述的用于无人机小目标监测的迁移学习方法, 其特征在于, 步骤S5的具体方法包括以下步骤:

S5-1、提取层级特征图的锚框和目标边框; 其中锚框和目标边框均为用两个顶点坐标表示的矩形边框;

S5-2、用两个边框的交集面积除以两个边框的并集面积, 得到两个边框的交并比;

S5-3、利用两个边框的交并比形成一个和层级特征图相同尺寸的矩阵, 即交并比分布图, 并计算该交并比分布图的平均值和方差;

S5-4、将交并比分布图的平均值和方差相加, 得到一个自适应阈值;

S5-5、判断层级特征图的特征点是否大于自适应阈值, 若是则判定得到匹配位置, 其值

为1,否则判定得到忽略点,其值为0;

S5-6、将判断结果组合成一个新的矩阵,即得到匹配点矩阵。

7.根据权利要求1所述的用于无人机小目标监测的迁移学习方法,其特征在于,步骤S6中采用特征点匹配损失函数作为代价函数计算L2范数损失的具体方法为:

根据公式:

$$loss_{trans} = \frac{\sum_{i=1}^M \sum_{j=1}^N (S(i, j) - T(i, j))^2 * mask(i, j)}{\sum_{i=1}^M \sum_{j=1}^N mask(i, j)}$$

得到归一化局部L2范数损失 $loss_{trans}$ ;其中, $S(i, j)$ 为对准后的源域网络的特征图上*i*行*j*列的值, $T(i, j)$ 为对准后的目标域网络的特征图上*i*行*j*列的值, $mask(i, j)$ 为匹配点矩阵, $M \times N$ 为特征图尺寸,且 $i \in (1, 2, \dots, M)$ ,  $j \in (1, 2, \dots, N)$ 。

## 用于无人机小目标监测的迁移学习方法

### 技术领域

[0001] 本发明涉及图像处理领域,具体涉及一种用于无人机小目标监测的迁移学习方法。

### 背景技术

[0002] 近年来,基于无人机遥感平台的目标检测,因无人机平台所具备的实时性高、覆盖面广、机动性强、成本低等特点,而在对地目标监视、城市巡逻等领域有广泛应用。然而,目前主流的目标检测算法在通用目标检测数据集如COCO和ImageNet等基准数据集上面达到了较高的精度,相比之下,在无人机数据集上的检测精度有较大的差距。这是因为COCO和ImageNet等数据集成像条件较好,包含的自然图像目标一般尺寸较大,具有辨别性的特征,且背景简单。相比之下,无人机图像目标更密集、尺寸更小,背景更加复杂,令小尺寸目标难以获得辨别性特征,极大增加了高精度目标识别的难度。

[0003] 为了提高小目标的检测精度,多尺度图像金字塔方法经常被用于增强小目标的特征表示。但该类方法的计算消耗巨大,图像长宽放大 $S$ 倍,计算量将增长 $S \times S$ 倍,在实际工程条件中很难应用。

### 发明内容

[0004] 针对现有技术中的上述不足,本发明提供一种用于无人机小目标监测的迁移学习方法解决了传统方法计算量大监测精度低的问题。

[0005] 为了达到上述发明目的,本发明采用的技术方案为:

[0006] 提供一种用于无人机小目标监测的迁移学习方法,其包括以下步骤:

[0007] S1、将无人机采集的开源对地图像数据作为训练样本;

[0008] S2、对以ResNet-50为骨干结构的RetinaNet网络进行初始化;

[0009] S3、通过初始化后的RetinaNet网络遍历训练样本,同时采用 $512 \times 512$ 的图像尺寸对训练样本进行迭代训练,确定目标域网络的初始参数,并形成目标域网络的特征图,采用 $1024 \times 1024$ 的图像尺寸对训练样本进行迭代训练,确定源域网络的最终参数,并形成源域网络的特征图;

[0010] S4、利用特征层级上采样对齐方式将目标域网络的特征图与源域网络的特征图进行对准;

[0011] S5、利用基于锚框交并比的自适应特征图匹配方式,将对准后的目标域网络与源域网络的层级特征图进行匹配,得到匹配点矩阵;

[0012] S6、根据对准后的目标域网络的特征图与源域网络的特征图和其匹配点矩阵,计算归一化局部L2范数损失,即源域网络到目标域网络对应关键特征点的残差;

[0013] S7、利用梯度反向传播进行最小化残差,缩小源域网络特征图与目标域网络特征图的差异,直到完成最小化残差,即完成源域到目标域的深度神经网络数据迁移训练学习,用于无人机小目标监测的迁移学习。

[0014] 进一步地：步骤S1中，采用的开源对地图像数据为VisDrone2018-DET数据集，训练样本为6471张。

[0015] 进一步地，步骤S2的具体方法包括以下子步骤：

[0016] S2-1、利用超过33万张图片的COCO数据对RetinaNet网络中的ResNet-50结构进行预训练；其中ResNet-50结构为卷积特征抽取网络；

[0017] S2-2、通过随机梯度下降方式对RetinaNet网络进行训练参数设置；其中初始学习率、动量和权重衰减系数分别设置为0.001、0.9、0.01。

[0018] 进一步地：步骤S3中遍历的次数为20轮，每次迭代的输入为4张图片压缩的张量，即每轮的迭代次数为训练样本的张数除以4；整个20轮遍历中，前1000次迭代学习率从0.001到0.01线性变化，除第6轮和第14轮遍历训练样本的学习率为前一轮的1/5，后续迭代的学习率与前一次相同。

[0019] 进一步地：步骤S4中特征层级上采样对齐方式为最近邻插值方式。

[0020] 进一步地，步骤S5的具体方法包括以下步骤：

[0021] S5-1、提取层级特征图的锚框和目标边框；其中锚框和目标边框均为用两个顶点坐标表示的矩形边框；

[0022] S5-2、用两个边框的交集面积除以两个边框的并集面积，得到两个边框的交并比；

[0023] S5-3、利用两个边框的交并比形成一个和层级特征图相同尺寸的矩阵，即交并比分布图，并计算该交并比分布图的平均值和方差；

[0024] S5-4、将交并比分布图的平均值和方差相加，得到一个自适应阈值；

[0025] S5-5、判断层级特征图的特征点是否大于自适应阈值，若是则判定得到匹配位置，其值为1，否则判定得到忽略点，其值为0；

[0026] S5-6、将判断结果组合成一个新的矩阵，即得到匹配点矩阵。

[0027] 进一步地，步骤S6中采用特征点匹配损失函数作为代价函数计算L2范数损失的具体方法为：

[0028] 根据公式：

$$[0029] \quad loss_{trans} = \frac{\sum_i^M \sum_j^N (S(i, j) - T(i, j))^2 * mask(i, j)}{\sum_i^M \sum_j^N mask(i, j)}$$

[0030] 得到归一化局部L2范数损失 $loss_{trans}$ ；其中， $S(i, j)$ 为对准后的源域网络的特征图上i行j列的值， $T(i, j)$ 为对准后的目标域网络的特征图上i行j列的值， $mask(i, j)$ 为匹配点矩阵， $M \times N$ 为特征图尺寸，且 $i \in (1, 2, \dots, M)$ ， $j \in (1, 2, \dots, N)$ 。

[0031] 本发明的有益效果为：将迁移学习应用于同样来源但尺寸不同的样本数据，通过大尺寸图像得到预训练模型，再利用迁移学习引导小尺度模型的训练，结合特征图的上采样对齐与自适应特征图匹配，从而有效提升小尺度模型的小目标检测精度，且避免了过多的计算量增长。直接采用采样方法获取多尺度图像，在保证良好效果的同时减低数据制备成本。利用大数据量公共图像数据集预训练，小数据量目标数据对模型微调的方法，减少了模型收敛所需的数据量，能够使模型更快地收敛，能够得到对目标数据集较好的检测结果。

## 附图说明

- [0032] 图1为本发明总流程图；  
[0033] 图2为本发明的算法架构示意图；  
[0034] 图3为本发明步骤S5的特征图匹配示意图。

## 具体实施方式

[0035] 下面对本发明的具体实施方式进行描述,以便于本技术领域的技术人员理解本发明,但应该清楚,本发明不限于具体实施方式的范围,对本技术领域的普通技术人员来讲,只要各种变化在所附的权利要求限定和确定的本发明的精神和范围内,这些变化是显而易见的,一切利用本发明构思的发明创造均在保护之列。

[0036] 如图1和图2所示,该用于无人机小目标监测的迁移学习方法包括以下步骤:

[0037] S1、将无人机采集的开源对地图像数据作为训练样本;

[0038] S2、对以ResNet-50为骨干结构的RetinaNet网络进行初始化;

[0039] S3、通过初始化后的RetinaNet网络遍历训练样本,同时采用 $512 \times 512$ 的图像尺寸对训练样本进行迭代训练,确定目标域网络的初始参数,并形成目标域网络的特征图,采用 $1024 \times 1024$ 的图像尺寸对训练样本进行迭代训练,确定源域网络的最终参数,并形成源域网络的特征图;

[0040] S4、利用特征层级上采样对齐方式将目标域网络的特征图与源域网络的特征图进行对准;

[0041] S5、利用基于锚框交并比的自适应特征图匹配方式,将对准后的目标域网络与源域网络的层级特征图进行匹配,得到匹配点矩阵;

[0042] S6、根据对准后的目标域网络的特征图与源域网络的特征图和其匹配点矩阵,计算归一化局部L2范数损失,即源域网络到目标域网络对应关键特征点的残差;

[0043] S7、利用梯度反向传播进行最小化残差,缩小源域网络特征图与目标域网络特征图的差异,直到完成最小化残差,即完成源域到目标域的深度神经网络数据迁移训练学习,用于无人机小目标监测的迁移学习。

[0044] 步骤S1中,采用的开源对地图像数据为VisDrone2018-DET数据集,训练样本为6471张。

[0045] 步骤S2的具体方法包括以下子步骤:

[0046] S2-1、利用超过33万张图片的COCO数据对RetinaNet网络中的ResNet-50结构进行预训练;其中ResNet-50结构为卷积特征抽取网络;

[0047] S2-2、通过随机梯度下降方式对RetinaNet网络进行训练参数设置;其中初始学习率、动量和权重衰减系数分别设置为0.001、0.9、0.01。

[0048] 步骤S3中遍历的次数为20轮,每次迭代的输入为4张图片压缩的张量,即每轮的迭代次数为训练样本的张数除以4;整个20轮遍历中,前1000次迭代学习率从0.001到0.01线性变化,除第6轮和第14轮遍历训练样本的学习率为前一轮的1/5,后续迭代的学习率与前一次相同。

[0049] 步骤S4中特征层级上采样对齐方式为最近邻插值方式。

[0050] 如图3所示,步骤S5的具体方法包括以下步骤:

[0051] S5-1、提取层级特征图的锚框和目标边框；其中锚框和目标边框均为用两个顶点坐标表示的矩形边框；

[0052] S5-2、用两个边框的交集面积除以两个边框的并集面积，得到两个边框的交并比；

[0053] S5-3、利用两个边框的交并比形成一个和层级特征图相同尺寸的矩阵，即交并比分布图，并计算该交并比分布图的平均值和方差；

[0054] S5-4、将交并比分布图的平均值和方差相加，得到一个自适应阈值；

[0055] S5-5、判断层级特征图的特征点是否大于自适应阈值，若是则判定得到匹配位置，其值为1，否则判定得到忽略点，其值为0；

[0056] S5-6、将判断结果组合成一个新的矩阵，即得到匹配点矩阵。

[0057] 图3中，真实标注即目标边框，IOU计算即交并比计算，Mean表示交并比分布图的平均值，std表示交并比分布图的方差， $\lambda_{\text{thresh}}$ 表示自适应阈值，以及match\_points表示匹配点矩阵。

[0058] 步骤S6中采用特征点匹配损失函数作为代价函数计算L2范数损失的具体方法为：

[0059] 根据公式：

$$[0060] \quad \text{loss}_{\text{trans}} = \frac{\sum_i^M \sum_j^N (S(i, j) - T(i, j))^2 * \text{mask}(i, j)}{\sum_i^M \sum_j^N \text{mask}(i, j)}$$

[0061] 得到归一化局部L2范数损失 $\text{loss}_{\text{trans}}$ ；其中， $S(i, j)$ 为对准后的源域网络的特征图上i行j列的值， $T(i, j)$ 为对准后的目标域网络的特征图上i行j列的值， $\text{mask}(i, j)$ 为匹配点矩阵， $M \times N$ 为特征图尺寸，且 $i \in (1, 2, \dots, M)$ ， $j \in (1, 2, \dots, N)$ 。

[0062] 本发明的一个具体实施例中，实验数据表明，利用大尺度图像来提升小目标检测精度的图像金字塔方法中，相比 $512 \times 512$ 尺寸图像训练的模型， $1024 \times 1024$ 图像训练的模型在小目标精度上提升了339.2%，但4倍上采样会使计算量增大到原来的4倍，即增长300%，这是因为图像金字塔方法在图像输入进行上采样，带来了巨大的计算量增长。相比之下，本发明的迁移训练方法对目标域特征图进行4倍上采样，对齐到源域网络，在只增加8.9%计算量的条件下，使小目标检测精度提升114.2%。本专利的精度-计算量增益比（小目标精度提升比例/计算量增长比例）为12.83，而图像金字塔方法增益比为1.13。

[0063] 本发明将迁移学习应用于同样来源但尺寸不同的样本数据，通过大尺寸图像得到预训练模型，再利用迁移学习引导小尺度模型的训练，结合特征图的上采样对齐与自适应特征图匹配，从而有效提升小尺度模型的小目标检测精度，且避免了过多的计算量增长。

[0064] 直接采用采样方法获取多尺度图像，在保证良好效果的同时减低数据制备成本。利用大数据量公共图像数据集预训练，小数据量目标数据对模型微调的方法，减少了模型收敛所需的数据量，能够使模型更快地收敛，能够得到对目标数据集较好的检测结果。

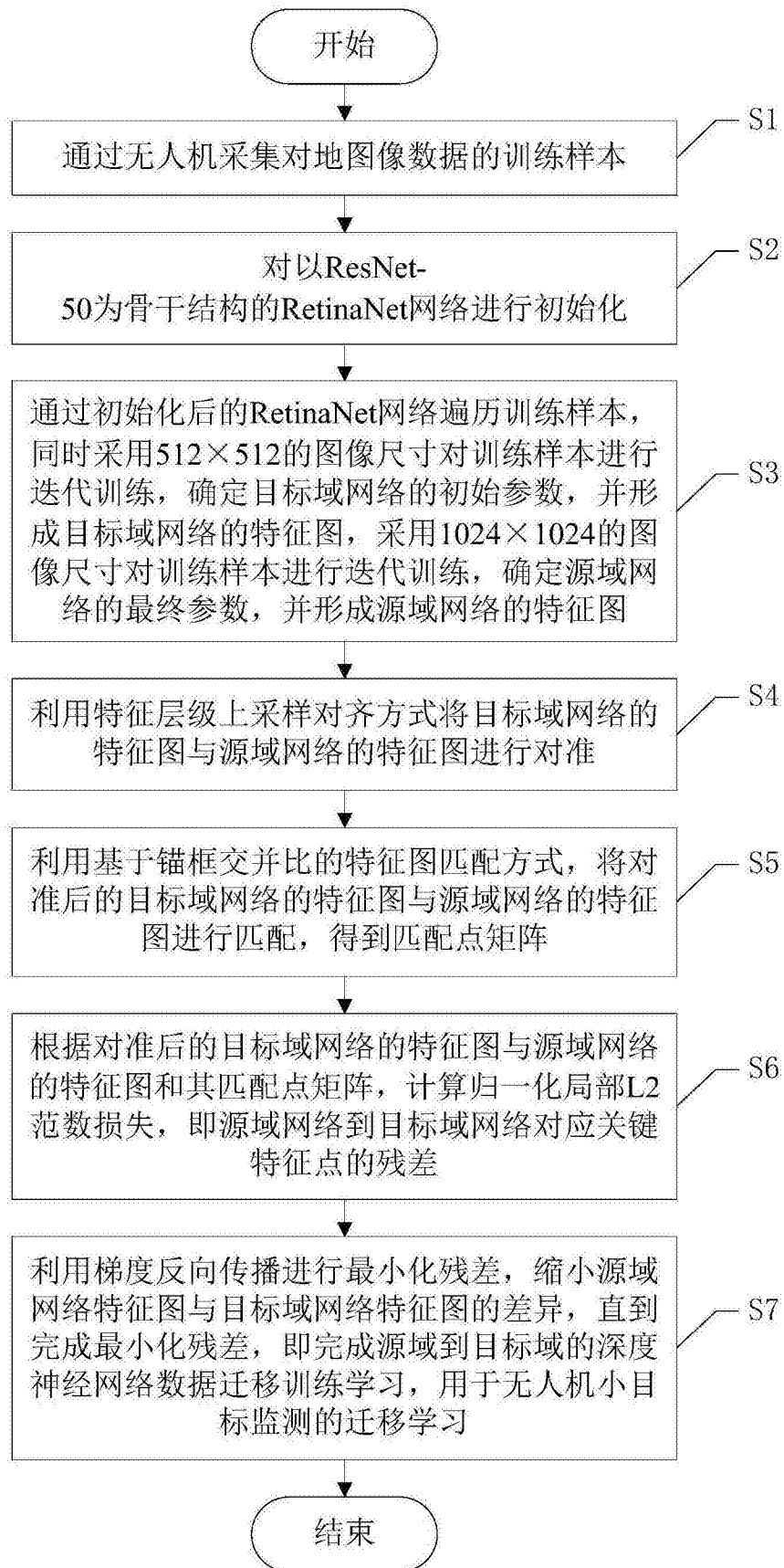


图1



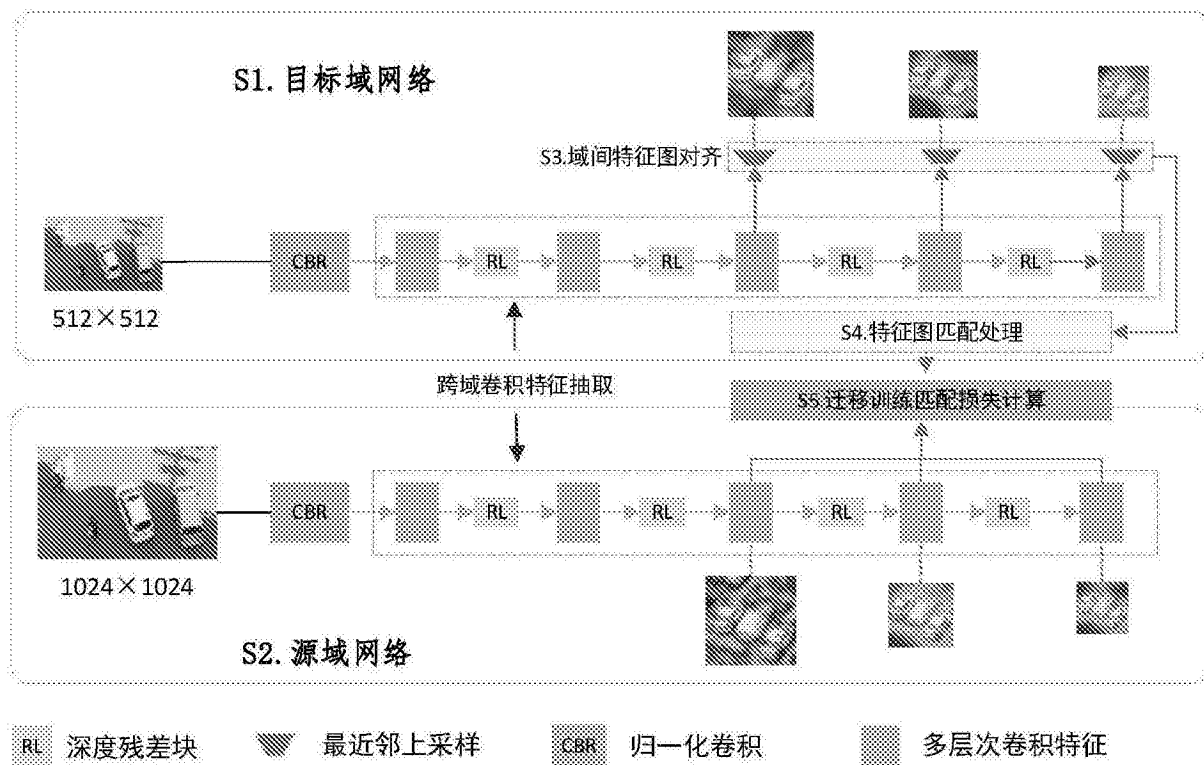


图2

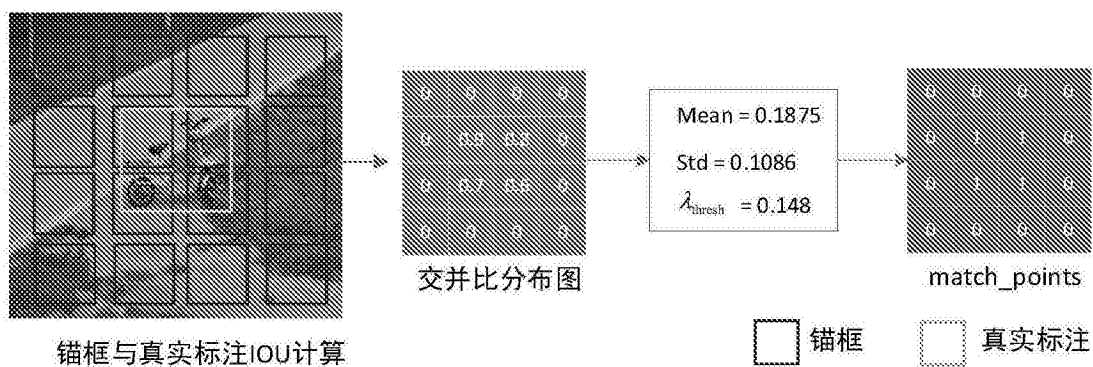


图3