



(12) 发明专利申请

(10) 申请公布号 CN 111768415 A

(43) 申请公布日 2020.10.13

(21) 申请号 202010542619.0

(22) 申请日 2020.06.15

(71) 申请人 哈尔滨工程大学

地址 150001 黑龙江省哈尔滨市南岗区南通大街145号哈尔滨工程大学科技处知识产权办公室

(72) 发明人 苏丽 孙雨鑫 苑守正

(51) Int.Cl.

G06T 7/11 (2017.01)

G06T 7/194 (2017.01)

G06K 9/62 (2006.01)

G06K 9/20 (2006.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

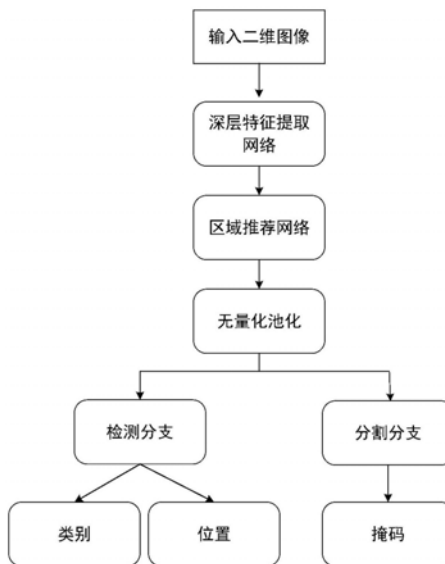
权利要求书1页 说明书5页 附图3页

(54) 发明名称

一种无量化池化的图像实例分割方法

(57) 摘要

本发明提供一种无量化池化的图像实例分割方法,包括如下步骤:S1:将任意大小二维图像输入深层特征提取网络获得多层特征图,经过区域推荐网络提取候选区域;S2:使用无量化池化层将上述具有不同的尺寸候选区域特征图池化到相同的大小;S3:输入两个检测分支预测每个候选区域的类别以及位置,同时并行连接掩码分支对每个候选区域进行前景和背景掩码分割,还原到原图尺寸。本发明解决了现有技术中对不同大小候选区域的特征图进行池化时像素空间信息缺失问题,使用无量化池化层,在不引入任何参数条件下,使得特征图像素与原图像素一一对应,保证了目标位置的准确性,进而提升了图像实例分割的准确度。



1. 一种无量化池化的图像实例分割方法,其特征是,包括如下步骤:

S1: 将任意大小二维图像输入深层特征提取网络获得多层特征图,经过区域推荐网络提取候选区域;构建深层特征提取网络,使用残差卷积网络和特征金字塔进行不同特征层融合,充分利用图像浅层位置信息和深层语义信息;区域推荐网络提取候选区域并区分前景和背景;

S2: 使用无量化池化层将上述具有不同的尺寸候选区域特征图池化到相同的大小;

S3: 输入两个检测分支预测每个候选区域的类别以及位置,同时并行连接掩码分支对每个候选区域进行前景和背景掩码分割,还原到原图尺寸。

2. 根据权利要求1所述的无量化池化的图像实例分割方法,其特征是,所述构建深层特征提取网络是使用101层残差卷积网络中4个阶段激活后的特征图和4层特征金字塔融合,将任意大小二维图像输入深层特征提取网络获得4层相同通道不同尺寸的特征图。

3. 根据权利要求1所述的无量化池化的图像实例分割方法,其特征是,所述区域推荐网络是卷积核为3\*3,通道数为256的卷积层。

4. 根据权利要求1所述的无量化池化的图像实例分割方法,其特征是,所述前景是有目标的位置,所述背景是没有目标的位置。

5. 根据权利要求1所述的无量化池化的图像实例分割方法,其特征是,所述步骤S2具体为使用无量化池化层将候选区域离散特征图插值到连续空间,直接在连续特征图上计算二重积分再除以面积得到均值。

6. 根据权利要求1所述的无量化池化的图像实例分割方法,其特征是,所述两个检测分支为多层全连接分类和回归网络。

7. 根据权利要求1所述的无量化池化的图像实例分割方法,其特征是,所述掩码分支为全卷积网络,包括4层卷积和1层反卷积。

## 一种无量化池化的图像实例分割方法

### 技术领域

[0001] 本发明涉及一种图像实例分割方法,尤其涉及一种无量化池化的图像实例分割方法,属于计算机视觉二维图像处理领域。

### 背景技术

[0002] 近年来,深度学习领域发展迅速直接推动了计算机视觉中图像处理领域进入了新的技术时代,图像实例分割作为基础研究问题,其技术可广泛应用于汽车自动驾驶,机器人控制和视频监控等领域。

[0003] 图像实例分割视为目标检测和语义分割的结合体,因为它需要正确分离图像中的所有实例,同时还要在像素级别对每个实例进行语义分割赋予不同的类别标签。比如:分割出一张图片中存在所有狗,还要具体表明不同狗,赋标为狗1,狗2。

[0004] 图像实例分割方法主要有两阶段和一阶段,其中Mask R-CNN是两阶段方法的经典之作,作为许多衍生应用的基线算法,继承了两阶段目标检测Faster R-CNN的思想,并扩展全卷积网络来预测目标掩码信息。实例分割挑战性在于处理图像过程中要很好地利用像素空间位置信息来区别图中属于同一种类不同实例的目标。Mask R-CNN中由区域推荐网络生成的不同候选区域位置是由模型回归得到的,通常是浮点数,在进行池化时为了获取浮点数位置的像素值使用双线性插值法,求一个浮点坐标的像素值需要选取固定4个采样点,即周围上下左右四个整数像素点,取值固定对不同图像不具有适应性,没有充分考虑候选区域对应特征图中每一点像素,间接存在量化取整问题,丢失了像素空间信息,存在像素偏差使得目标位置不准确,根据此位置无法准确分割目标掩码,进而影响后续实例分割准确度。

[0005] 通过对现有方法的探索,并未发现类似专利。因此针对现有技术中池化过程存在像素空间信息缺失的问题,提出一种无量化池化的图像实例分割对提升分割准确度是有一定意义的。

### 发明内容

[0006] 本发明的目的在于提供一种无量化池化的图像实例分割方法,特点在于对不同大小的候选区域特征图进行池化时保留完整的像素空间信息,保证了目标位置的准确性,进而提升实例分割准确度,特别是图像小目标。

[0007] 本发明的目的是这样实现的:

[0008] 一种无量化池化的图像实例分割方法,包括如下步骤:

[0009] S1:将任意大小二维图像输入深层特征提取网络获得多层特征图,经过区域推荐网络提取候选区域;构建深层特征提取网络,使用残差卷积网络和特征金字塔进行不同特征层融合,充分利用图像浅层位置信息和深层语义信息;区域推荐网络提取候选区域并区分前景和背景;

[0010] S2:使用无量化池化层将上述具有不同的尺寸候选区域特征图池化到相同的大小;

[0011] S3:输入两个检测分支预测每个候选区域的类别以及位置,同时并行连接掩码分支对每个候选区域进行前景和背景掩码分割,还原到原图尺寸;

[0012] 本发明还包括这样一些特征:

[0013] 所述构建深层特征提取网络是使用101层残差卷积网络中4个阶段激活后的特征图和4层特征金字塔融合,将任意大小二维图像输入深层特征提取网络获得4层相同通道不同尺寸的特征图;

[0014] 所述区域推荐网络是卷积核为3\*3,通道数为256的卷积层;

[0015] 所述前景是有目标的位置,所述背景是没有目标的位置;

[0016] 所述步骤S2具体为使用无量化池化层将候选区域离散特征图插值到连续空间,直接在连续特征图上计算二重积分再除以面积得到均值;

[0017] 所述两个检测分支为多层全连接分类和回归网络;

[0018] 所述掩码分支为全卷积网络,包括4层卷积和1层反卷积。

[0019] 与现有技术相比,本发明的有益效果是:

[0020] 1、本发明解决了不同大小候选区域的特征图进行池化时像素空间信息缺失问题,更好保证目标的位置准确性;使用无量化池化层,在不引入任何参数条件下,使得特征图像素与原图像素一一对应,保证了目标位置的准确性,进而提升了图像实例分割的准确度;

[0021] 2、本发明全面提升了普通场景下的二维图像实例分割的准确度,特别是小目标的提升。

## 附图说明

[0022] 图1是本发明一种无量化池化的图像实例分割方法的流程图;

[0023] 图2是本发明一种无量化池化的图像实例分割方法的深层特征提取网络示意图;

[0024] 图3是本发明一种无量化池化的图像实例分割方法的ROIAlign示意图;

[0025] 图4是本发明一种无量化池化的图像实例分割方法的无量化池化示意图。

## 具体实施方式

[0026] 下面结合附图与具体实施方式对本发明作进一步详细描述。

[0027] 本发明内容步骤如下:

[0028] S1:将任意大小二维图像输入深层特征提取网络获得多层特征图,经过区域推荐网络提取候选区域。

[0029] 构建深层特征提取网络,使用残差卷积网络和特征金字塔进行不同特征层融合,充分利用图像浅层位置信息和深层语义信息。区域推荐网络提取候选区域,经分类与回归分支进行前景与后景的二分类(不区分具体所属类别)和候选区域位置的微调;

[0030] S2:使用无量化池化层将上述具有不同尺寸的候选区域特征图池化到相同的大小。与原来的Mask R-CNN池化方法相比,无量化池化层直接计算基于连续特征图的二重积分再取平均值,充分保留了像素空间信息。它不需要预先定义采样点的数目,避免了坐标的任何量化,并且在候选区域坐标上具有连续的梯度。

[0031] 将候选区域按坐标位置从原图像上映射到特征图上,使用插值法将离散的特征图映射到一个连续空间。在连续特征图上对其像素进行二重积分取平均值作为池化后的像素

值。

$$[0032] \quad P(g, \mathcal{F}) = \frac{\int_{y_1}^{y_2} \int_{x_1}^{x_2} \mathcal{F}(x, y) dx dy}{(x_2 - x_1) \times (y_2 - y_1)}$$

[0033] S3:输入两个检测分支预测每个候选区域的类别以及位置,同时并行掩码分支连接全卷积网络对每个候选区域进行前景和背景掩码分割。

[0034] 本发明所述方法使用开源COCO数据集中135000张普通场景图片和对应真实标注文件作为训练集,5000张图片和对应真实标注文件作为验证集和COCO分割指标进行实例分割准确度评价。

[0035] 参照图一,介绍本专利一种无量化池化的图像实例分割方法的具体实施方式,具体包括

[0036] S1:将任意大小二维图像输入深层特征提取网络获得多层特征图,经过区域推荐网络提取候选区域。

[0037] S11:所述二维图像可为三通道RGB彩色图像。

[0038] S12:构建深层特征提取网络,参照图二,使用101层残差卷积网络和4层特征金字塔进行高低层不同特征层融合,充分利用图像浅层位置信息和深层语义信息。残差卷积网络4个阶段激活特征图为 $\{C_2, C_3, C_4, C_5\}$ , (其中 $C_1$ 层参数量大占用内存没有被选用)融入特征金字塔后特征图为 $\{P_2, P_3, P_4, P_5\}$ ,所有特征图通道数均为256。具体如 $C_5$ 经 $1*1$ 卷积操作降维为 $P_5$ ,对 $P_5$ 进行2倍上采样与经 $1*1$ 卷积操作后的 $C_4$ 进行相加操作融合生成 $P_4$ ,其他特征层操作同理。

[0039] S13:经过区域推荐网络细化特征生成候选区域。区域推荐网络是目标检测任务中预处理图像产生候选区域常用的方法,通常在特征提取网络后连接卷积核为 $3*3$ ,通道数为256的卷积层。并行连接两个 $1*1$ 卷积形成分类与回归分支,进行候选区域前景与后景的二分类(不区分具体所属类别)和第一次候选区域位置的微调。

[0040] S2:参照图三,Mask RCNN中为了将候选区域对应的特征图池化到同一个尺度下送入后续全连接分类与回归网络中使用了ROIAlign算法,简要方法包括:

[0041] S21:使用双线性插值法获得坐标为浮点数的像素点对应图像像素值来补充出候选区域边界上的缺失坐标点,从而将候选区域池化过程转化为一个连续的操作,过程可参照图二。

[0042] 涉及到双线性插值法,为求一个点的像素需要固定采样周围四个像素点,取值固定不具有适应性,没有充分考虑候选区域对应特征图中每一点像素,存在坐标量化问题丢失了目标的空间位置信息,严重影响了后续目标的定位与分割。

[0043] S22:本发明中无量化池化层是对相应特征区域先求二重积分再取平均值,从而避免了坐标的任何量化,并且在边界框坐标上具有连续的渐变性,使得特征图像素与原图片像素一一对应,保证了目标位置的准确性。配合参照图四,具体算法包括:

[0044] 首先将候选区域按坐标位置从原图像上映射到特征图上,根据区域裁剪特征图,此时为离散特征图,为了保留更多细节信息,需对候选区域的特征图 $f$ 进行插值处理,将离散的特征图映射到一个连续空间得到连续特征图 $\mathcal{F}$ ,公式如下:

$$[0045] \quad \mathcal{F}(x, y) = \sum_{i, j} C(x, y, i, j) * f(i, j)$$

$$[0046] \quad C(x, y, i, j) = \max(0, 1 - |x - i|) * \max(0, 1 - |y - j|)$$

[0047] 其中  $(i, j)$  为离散特征图上位置坐标,  $f(i, j)$  表示对应坐标的像素值,  $(x, y)$  为连续特征图上位置坐标,  $\mathcal{F}(x, y)$  表示对应坐标的像素值,  $C(x, y, i, j)$  是插值系数。

[0048] 假设候选区域对应特征图尺寸为  $(W, H)$ , 经无量化池化后的尺寸大小为  $S * S$ , 则需要在每一小格  $g = (W/S, H/S)$  中选取一个像素值, 组成  $S * S$  大小的像素图。如果取  $(x_1, y_1, x_2, y_2)$  为候选区域某一小格坐标, 其中  $(x_1, y_1)$  为小格左上角坐标,  $(x_2, y_2)$  为小格右上角坐标, 这里  $(x, y)$  均为插值后的连续坐标, 无量化池化层为了充分利用特征图上所有像素信息, 对每一个小格  $g$  的像素值先进行二重积分再除以小格面积取平均值得到代表这个小格的像素值, 其他同理, 最后所有小格组合得到连续的特征图。公式如下:

$$[0049] \quad P(g, \mathcal{F}) = \frac{\int_{y_1}^{y_2} \int_{x_1}^{x_2} \mathcal{F}(x, y) dx dy}{(x_2 - x_1) \times (y_2 - y_1)}$$

[0050] 分析上式  $P(g, \mathcal{F})$  是连续可微的, 故候选区域可以参与神经网络的反向传播, 其中对  $x_1$  的偏导数计算如下:

$$[0051] \quad \frac{\partial P(g, \mathcal{F})}{\partial x_1} = \frac{P(g, \mathcal{F})}{x_2 - x_1} - \frac{\int_{y_1}^{y_2} \mathcal{F}(x, y) dy}{(x_2 - x_1) \times (y_2 - y_1)}$$

[0052]  $P(g, \mathcal{F})$  为连续函数, 其他坐标偏导数同理可求。

[0053] 通过上述方法改进, 在引入任何参数情况下可以实现候选区域坐标连续化, 考虑周围所有像素更好地保留像素空间信息, 使得特征图像素与原图片像素一一对应, 保证了目标位置的准确性, 特别是图像中的小目标 (比如, 同样是 0.5 个像素点的偏差, 对于较大的目标而言显得微不足道, 但是对于小目标, 误差的影响就要高很多)。

[0054] S3: 经无量化池化后具有同样大小的候选区域特征图输入两个分类器分支即两个多层全连接分类网络来预测每个候选区域的类别以及位置, 掩码分支与候选区域共享特征层经过无量化池化层, 后连接四层全卷积网络实现对每个可能有目标的候选区域进行前景和背景掩码分割, 最后得到图像中每个实例的类别种类, 位置坐标和分割掩码。

[0055] 本发明中的方法在开源 COCO 数据集上做了对比实验, 可以有效提升图像实例分割准确度。训练集为 135000 张普通场景图片, 验证集为 5000 张图片。

[0056] 准确度评价使用 COCO 数据集常用评价指标 AP (取所有 IOU 阈值平均值),  $AP_{50}$ ,  $AP_{75}$ ,  $AP_s$ ,  $AP_M$ ,  $AP_L$  在验证集上进行实验结果如下。

算法	深层提取网络	AP	$AP_{50}$	$AP_{75}$	$AP_s$	$AP_M$	$AP_L$
[0057] MaskRCNN	ResNet-101-C4	33.10	54.98	34.82	12.11	35.62	51.11
MaskRCNN	ResNet-101-FPN	36.15	58.10	38.33	16.41	38.91	53.37
本发明	ResNet-101-FPN	36.32	58.57	38.60	17.13	39.26	54.11

[0058] 可见本发明在同等设置下实例分割准确度高于 MaskRCNN, 特别是在图像小目标。

[0059] 综上所述: 本专利公开了一种无量化池化的图像实例分割方法, 主要用于普通场

景下二维图像的目标实例分割问题。主要步骤有：(1) 将任意大小二维图像输入深层特征提取网络获得多层特征图，经过区域推荐网络提取候选区域，(2) 使用无量化池化层对候选区域的特征图进行池化，(3) 输入两个检测分支预测每个候选区域的类别以及位置，同时并行连接掩码分支对每个候选区域进行前景和背景掩码分割。本发明解决了现有技术中对不同大小候选区域的特征图进行池化时像素空间信息缺失问题，使用无量化池化层，在不引入任何参数条件下，使得特征图像素与原图像像素一一对应，保证了目标位置的准确性，进而提升了图像实例分割的准确度。

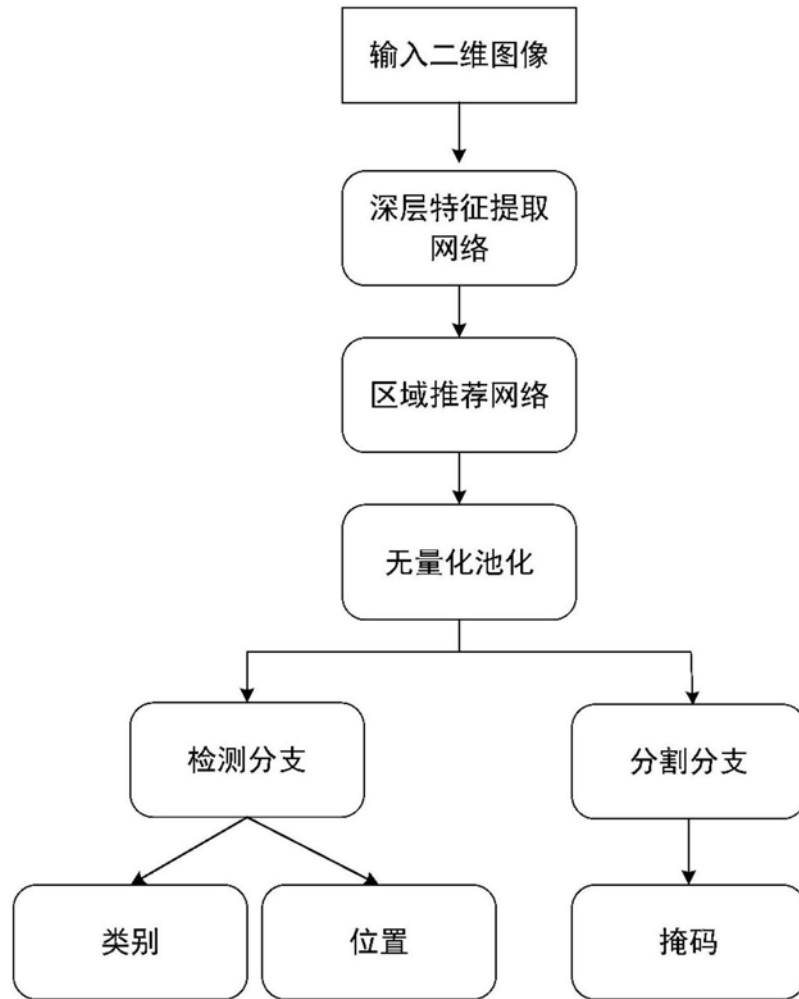


图1



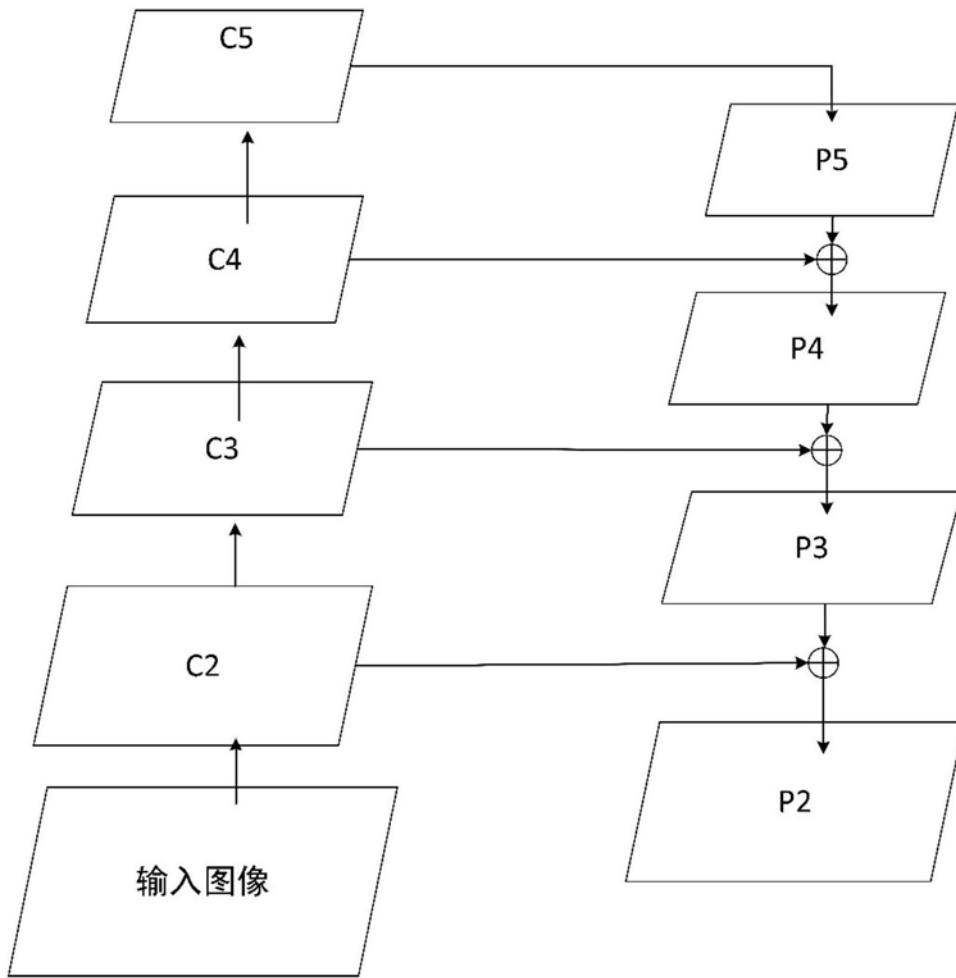


图2

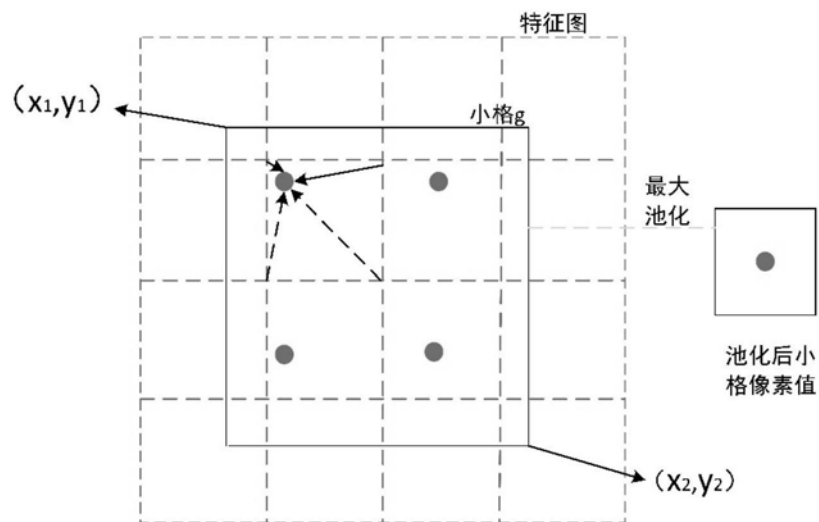


图3

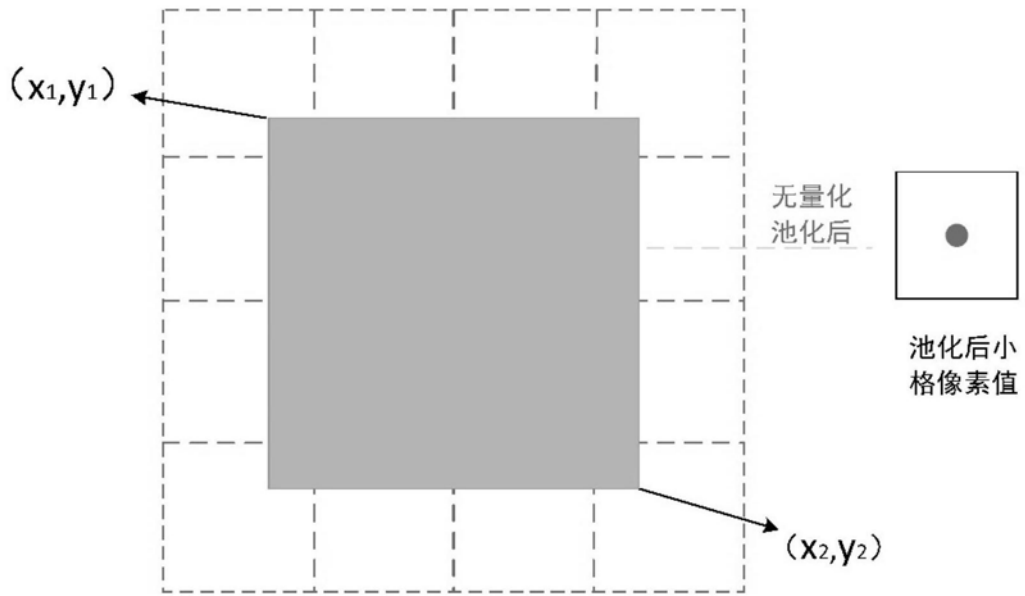


图4