# (19) 中华人民共和国国家知识产权局



# (12) 发明专利申请



(10) 申请公布号 CN 112084868 A (43) 申请公布日 2020. 12. 15

(21) 申请号 202010794525.2

(22)申请日 2020.08.10

(71) 申请人 北京航空航天大学 地址 100191 北京市海淀区学院路37号

(72) 发明人 刘庆杰 高广帅 王蕴红

(74) 专利代理机构 北京航智知识产权代理事务 所(普通合伙) 11668

代理人 黄川 史继颖

(51) Int.CI.

GO6K 9/00 (2006.01)

G06T 7/60 (2017.01)

GO6N 3/04 (2006.01)

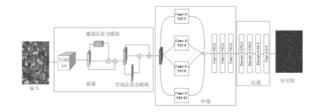
权利要求书1页 说明书7页 附图4页

## (54) 发明名称

一种基于注意力机制的遥感图像中目标计 数方法

#### (57) 摘要

本发明公开了一种基于注意力机制的遥感 图像中目标计数方法,在VGG16的基础上融合了 注意力机制,尺度金字塔和可变形卷积的特性, 由三个级联的阶段组成:前端网络的特征提取, 中端网络的尺度融合以及后端网络的密度图的 生成。通过本发明的技术方案,能够很好地解决 遥感图像中密集目标计数任务中存在的目标尺 度多样性、复杂混乱的背景干扰以及目标方向任 意性的问题。



- 1.一种基于注意力机制的遥感图像中目标计数方法,其特征在于,包括在VGG16网络结构基础上,对输入图像进行以下三个级联阶段的处理:
  - S1:前端网络的特征提取;

对输入图像,取VGG16网络结构的前10层操作,随后融合卷积块注意力模块,即将通道注意力模块和空间注意力模块连接的操作,用来编码特征图通道及像素位置之间的相关性;

S2:中端网络的多尺度融合;

引入尺度金字塔模块,级联膨胀率分别为2,4,8,12的空洞卷积,捕获更多的多尺度信息和细节信息;

S3:后端网络的密度图的生成;

采用三层卷积核为 $3\times3$ 的可变形卷积,每层之后都有一层修正线性单元ReLU激活函数,最后添加一个 $1\times1$ 的卷积层用来生成密度图:

S4: 将步骤S3的密度图的所有像素求和得到最终的目标数量。

# 一种基于注意力机制的遥感图像中目标计数方法

#### 技术领域

[0001] 本发明属于遥感图像技术领域,尤其涉及一种基于注意力机制的遥感图像中目标计数方法。

# 背景技术

[0002] 近几十年来,随着国家安防和城市规划等各方面的需要,估计复杂场景中目标的数目越来越受到关注。因此,已经有很多的工作用到各个领域的目标计数上,例如监控视频中的人群计数,显微镜下的细胞计数,生态研究的动物计数,车辆计数和环境研究中的目标计数。

[0003] 尽管目标计数在各个领域进步很大,但是在遥感领域中还很少涉及。除了一些很少的场景,例如棕榈树或者橄榄树的计数,无人机拍摄图片中的车辆计数。但是,遥感图像中主要的地物目标,比如建筑物,船只等,没有受到过多的关注。因此,对这些目标进行计数能够有很多的实际意义,比如城市规划,环境控制,数字城市模型构建和对灾难的应对方案规划等。

[0004] 相比于其他领域的目标计数,遥感图像的目标计数存在以下几个挑战:1)尺度多样性:遥感图像中的目标尺度变化不一,例如同一张图片中的尺寸从几个像素到千级像素不等;2)背景复杂多样:遥感图像中通常同时存在多种地物目标,特别当目标的尺寸很小的情况下,复杂混乱的背景干扰大大限制了目标的检测和计数;3)方向任意性:不同于自然场景图片中的目标,例如行人是直立的,由于星载或机载传感器俯拍视角的原因,遥感图像中的目标具有任意的方向。

[0005] VGG名称来源于牛津大学科学工程系的视频几何组(Visual Geometry Group),其发布了一系列以VGG开头的卷积网络模型,可以应用在人脸识别、图像分类等方面,分别从VGG11~VGG19。VGG研究卷积网络深度的初衷是为了探寻卷积网络深度是如何影响大规模图像分类与识别的精度和准确率的,VGG在加深网络层数的同时为了避免参数过多,在所有层都采用3x3的小卷积核,卷积层步长被设置为1。VGG的输入被设置为224x244大小的RGB图像,在训练集图像上对所有图像计算RGB均值,然后把图像作为输入传入VGG卷积网络,使用3x3或者1x1的filter,卷积步长被固定1。VGG全连接层有3层,根据卷积层+全连接层总数目的不同可以从VGG11~VGG19,最少的VGG11有8个卷积层与3个全连接层,最多的VGG19有16个卷积层+3个全连接层,此外VGG网络并不是在每个卷积层后面跟上一个池化层,总共有5个池化层,分布在不同的卷积层之下。VGG16最初应用在图像分类任务上,因其简洁性和实用性等特点,迅速成为当时最流行的卷积神经网络模型,现在也常被应用于各种计算机视觉任务中。

## 发明内容

[0006] 为了解决遥感图像中密集目标计数任务中存在的数据集稀缺、目标尺度多样性、复杂混乱的背景干扰以及目标方向任意性的问题,本发明提出一种基于注意力机制的遥感

图像中目标计数方法,在VGG16网络结构基础上融合了注意力机制(attention mechanism),尺度金字塔(scale pyramid)和可变形卷积(deformable convolution)的特性,简称为ASPDNet,由三个级联的阶段组成:前端网络的特征提取,中端网络的尺度融合以及后端网络的密度图的生成。本发明的具体技术方案如下:

[0007] 一种基于注意力机制的遥感图像中目标计数方法,其特征在于,包括在VGG16网络结构基础上,对输入图像进行以下三个级联阶段的处理:

[0008] S1:前端网络的特征提取;

[0009] 对输入图像,取VGG16网络结构的前10层操作,随后融合卷积块注意力模块,即将通道注意力模块和空间注意力模块连接的操作,用来编码特征图通道及像素位置之间的相关性:

[0010] S2:中端网络的多尺度融合:

[0011] 引入尺度金字塔模块,级联膨胀率分别为2,4,8,12的空洞卷积,捕获更多的多尺度信息和细节信息;

[0012] S3:后端网络的密度图的生成;

[0013] 采用三层卷积核为3×3的可变形卷积,每层之后都有一层修正线性单元ReLU激活函数,最后添加一个1×1的卷积层用来生成密度图:

[0014] S4:将步骤S3的密度图的所有像素求和得到最终的目标数量。

[0015] 本发明的有益效果在于:

[0016] 1.前端网络是以VGG16网络结构的前10层为主干网络,随后添加注意力模块,能够突出更感兴趣的目标区域突出出来,复杂的背景区域被抑制,能够很好地解决遥感图像中复杂混乱的背景干扰的问题。

[0017] 2.在网络的中端引入尺度金字塔模块,在不增加参数量的情况下,捕获对应不同感受野的多尺度信息,从而能够很好的解决尺度多样性的问题。

[0018] 3. 后端网络中采用三层可变形卷积,卷积中学习的偏置能够很好地覆盖目标,从而能够很好的解决遥感图像中目标任意性的问题。

#### 附图说明

[0019] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,通过参考附图会更加清楚的理解本发明的特征和优点,附图是示意性的而不应理解为对本发明进行任何限制,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,可以根据这些附图获得其他的附图。其中:

[0020] 图1是本发明的网络结构流程图:

[0021] 图2(a)是通道注意力模块结构图:

[0022] 图2(b)是空间注意力模块结构图:

[0023] 图3是尺度金字塔模块示意图:

[0024] 图4是可变形卷积原理图:

[0025] 图5(a)是标准卷积位置采样可视图:

[0026] 图5(b)是可变形卷积位置采样可视图;

[0027] 图6(a)是建筑物图片;

- [0028] 图6(b)是建筑物真实的密度图及计数结果;
- [0029] 图6(c)是本发明的方法得到的建筑物密度图及计数结果;
- [0030] 图7(a) 是小车的图片;
- [0031] 图7(b)是小车真实的密度图及计数结果;
- [0032] 图7(c)是本发明的方法得到小车的密度图及计数结果;
- [0033] 图8(a)是大型车辆的图片;
- [0034] 图8(b)是大型车辆真实的密度图及技术结果;
- [0035] 图8(c)是本发明的方法得到的大型车辆的密度图及计数结果;
- [0036] 图9(a) 是船只的图片;
- [0037] 图9(b) 是船只真实的密度图及计数结果;
- [0038] 图9(c)是本发明的方法得到的船只的密度图及计数结果。

## 具体实施方式

[0039] 为了能够更清楚地理解本发明的上述目的、特征和优点,下面结合附图和具体实施方式对本发明进行进一步的详细描述。需要说明的是,在不冲突的情况下,本发明的实施例及实施例中的特征可以相互组合。

[0040] 在下面的描述中阐述了很多具体细节以便于充分理解本发明,但是,本发明还可以采用其他不同于在此描述的其他方式来实施,因此,本发明的保护范围并不受下面公开的具体实施例的限制。

[0041] 本发明旨在精确估计遥感图像中密集目标的数目,例如连片密集的房屋,停泊在港口的船只,停在停车场的小车或者大型货车等。

[0042] 针对要解决的技术问题,本发明构建了一个基于密度估计的目标计数的卷积神经网络框架,在VGG16网络结构基础上融合了注意力机制(通道注意力和空间注意力模块连接的形式)模块,尺度金字塔模块和可变形卷积模块等技术,具体地,主要由三个级联的阶段组成:前端网络的特征提取,中端网络的尺度融合以及后端网络的密度图的生成,最后,对密度图中所有像素进行求和,即可计算出遥感图像中目标的数目。具体的网络流程图如图1所示。

[0043] 前端网络是以VGG16网络结构的前10层为主干网络,添加注意力模块,充分考虑特征图通道之间和像素位置之间的相关性,提取丰富的语义信息和上下文信息,从而能够突出更感兴趣的目标区域突出出来,复杂的背景区域被抑制,能够很好地解决遥感图像中复杂混乱的背景干扰的问题。

[0044] 由于在网络中应用到三层最大池化层,图像的分辨率会降低到原始图像的1/64,为了扩大特征图的感受野,在网络的中端引入尺度金字塔模块(scale pyramid module, SPM),即将四个平行的不同膨胀因子的空洞卷积连接起来,SPM的作用是在不增加参数量的情况下,捕获对应不同感受野的多尺度信息,从而能够很好的解决尺度多样性的问题。

[0045] 后端网络中采用三层可变形卷积(deformable convolution),可变形卷积操作就是在原始标准的卷积上增加一个可学习的偏置,得益于可变形卷积中的自适应位置采样技术,卷积中学习的偏置能够很好地覆盖目标,从而能够很好的解决遥感图像中目标任意性的问题,在网络的最后一层用一个1×1的卷积层,从而生成密度图,最后将密度图的所有像

素进行求和从而获得目标的数量。具体地,

[0046] 1.前端网络的特征提取

[0047] 给定一幅任意尺寸的遥感图片,以VGG16网络结构作为主干网络,取VGG16网络结构的前10层操作,然后添加卷积块注意力模块,是一个连接通道注意力模块(channel attention module,CAM)和空间注意力的模块(spatial attention module,SAM)的操作,用来编码特征图通道及像素位置之间的相关性,以此来采集更为主要的特征的信息,从而达到突出目标并且抑制复杂混乱背景的目的。

[0048] 通道注意力模块:在密集的场景中,前景目标与背景的纹理很相似,使得计数很困难,融入通道注意力模块能够减轻这个问题,通道注意力模块的结构体系如图2 (a) 所示。具体地,对于任意一个中间层的特征图表示为 $F \in \mathbb{R}^{C \times H \times W}$ ,其中,见表示实数空间,C表示特征图的通道,H和W分别表示特征图的高和宽,首先,对特征图执行一个1×1的卷积操作,然后通过变形和转置获得两个特征图 $C_1$ 和 $C_2$ ;接下来,将 $C_1$ 和 $C_2$ 相乘并执行归一化指数 (softmax)操作,获得尺寸为 $C \times C$ 的通道注意图 $C_a$ 。特别地,这个过程表示为:

$$C_a^{ji} = \frac{\exp\left(C_1^i \cdot C_2^j\right)}{\sum_{i=1}^c \exp\left(C_1^i \cdot C_2^j\right)}$$
,其中, $C_a^{ji}$ 表示通道注意图上第i个通道对第j个通道的影响, $C_1^i$ 为对

原始特征图先经过 $1\times 1$ 卷积后再变形为尺寸为 $C\times HW$ 的特征图的第i个通道, $C_2^j$ 为原始特征图先经过 $1\times 1$ 卷积后再变形转置为尺寸为 $HW\times C$ 的特征图的第j个通道,最后经过通道注意力模块加权的尺寸为 $C\times H\times W$ 的特征图计算: $C_{final}^j = \lambda \sum_{i=1}^{C} \left(C_a^{ji}\cdot C_3^i\right) + F^j$ ,其中, $\lambda$ 是一个可学

习的参数,可以通过 $1 \times 1$ 的卷积操作学习得到, $C_{final}^{i}$ 为最后经过通道注意力模块加权后的特征图的第 $_{1}^{i}$ 个通道, $C_{3}^{i}$ 为对原始特征图先经过 $1 \times 1$ 卷积后再变形为尺寸为 $C \times HW$ 的特征图的第 $_{1}^{i}$ 个通道, $F_{3}^{i}$ 为原始的特征图的第 $_{1}^{i}$ 个通道。

[0049] 空间注意力模块:考虑到特征图在不同的像素位置上有不同的密度,进一步对空间维度上长范围的依赖性进行编码,从而对空间位置上的特征信息编码很有效。空间注意力模块与上述的通道注意力模块相似,具体的网络架构如图2(b)所示。但是,两者有不同之处:1)相比于通道注意力模块只有一个1×1的卷积层,空间注意力模块需要三个;2)相比于通道注意 图 C 。的 C 寸 为 C × C ,空 间 注 意 图 C 。的 C 寸 为 C × C ,空 间 注 意 图 C 。的 C 寸 为 C × C ,空 间 注 意 图 C 。

$$S_a^{kl} = \frac{\exp\left(S_1^k \cdot S_2^l\right)}{\sum_{k=1}^{HW} \exp\left(S_1^k \cdot S_2^l\right)}$$
,其中, $S_a^{lk}$ 表示第k个位置对第1个位置的影响, $S_1^k$ 为原始特征图先经

过 $1 \times 1$ 卷积后再变形为尺寸为 $C \times HW$ 的特征图的第k个位置, $S_2^l$ 为原始特征图先经过 $1 \times 1$ 卷积后再变形转置为尺寸为 $HW \times C$ 的特征图的第1个位置,最后经过空间注意力模块加权的尺寸为 $C \times H \times W$ 的特征图计算如下: $S_{final}^l = \lambda \sum_{k=1}^{HW} \left( S_a^k \cdot S_3^k \right) + F^l$ ,其中, $\mu$ 是一个可学习的参数,可

以通过 $1 \times 1$ 的卷积操作学习得到, $S_{final}^{l}$ 为最终经过空间注意力模块加权的特征图的第1个位置, $S_{3}^{k}$ 为原始特征图先经过 $1 \times 1$ 卷积后再变形为尺寸为 $C \times HW$ 的特征图的第k个位置, $F^{1}$ 为原始特征图的第1个位置。

[0050] 2.中端网络的多尺度融合

[0051] 因为在前端网络中存在三层的池化层操作,输出特征图的尺寸是原始尺寸的1/64。为了在保持分辨率不变的同时增大特征图的感受野,引入如图3所示的尺度金字塔模块,即级联几个不同膨胀率的空洞卷积的操作。空洞卷积的作用就是在不增加参数量和复杂度的前提下,增大了特征图的感受野,而不同的膨胀率对应不同大小的感受野。在本发明的方法中,设定级联的空洞卷积数目为4,采用的膨胀率分别为2,4,8,12,通过尺度金字塔模块操作,能够捕获更多的多尺度信息和细节信息,从而提升了模型对尺度变化的鲁棒性。

[0052] 3.后端网络的密度图的生成

[0053] 在后端网络中,采用三层可变形卷积操作来解决遥感图像中目标方向任意性的问题,最后添加一个1×1的卷积层用来生成密度图。

[0054] 可变形卷积 (deformable convolution) 操作,相比于标准卷积,就是在特征图的感受野的每个像素点上增加一个大小可以学习的偏置。学习这个偏置的作用就是无论目标的形状如何变化,卷积层都能够覆盖到整个目标。可变形卷积的原理图和具体的位置采样的可视图由图4、图5 (a) 和图5 (b) 所示。

[0055] 对于一个标准的卷积,给定一个采样点的位置 $p_m$ ,对于一个卷积核为 $3 \times 3$ ,膨胀率为1的卷积, $\mathbf{p}_m \in \mathcal{M} = \{(-1,-1),(0,-1),...,(1,0),(1,1)\}$ , $\mathcal{M}$ 为规则采样点的集合,位置p的输出特征图  $y(\mathbf{p}) = \sum_{m=1}^{\mathcal{M}} \mathbf{w} \left( \mathbf{p}_m \right) \cdot \mathbf{x} (\mathbf{p} + \mathbf{p}_m)$ ,其中, $\mathbf{w}$ 表示加权的参数, $\mathbf{x}$ 表示输入的特征图,

 $m=1,\cdots,M, M=|\mathcal{M}|$ ,表示第m个采样点,M为采样点的总数。相比于标准卷积,可变形卷积就是在此基础上增加一个可通过训练优化得到一个可以自适应学习的偏置  $\Delta p_m$ ,对于可变形

卷积,特征图y(p) =  $\sum_{m=1}^{M} w(\mathbf{p}_m) \cdot x(\mathbf{p} + \mathbf{p}_m + \Delta \mathbf{p}_m)$ ,具体地,采用三层卷积核为3×3的可变形卷

积,每层之后都有一层修正线性单元(Rectified Linear Unit,ReLU)激活函数。通过这种可变形卷积中的动态采样的策略,遥感图像中由于俯视视角导致的目标方向任意性能够得到很好的解决。在网络的最后,加上一层1×1的卷积层用来生成密度图。将密度图的所有像素求和就可以得到最终的目标数量。

[0056] 本发明的方法是从输入图片估计出密度图来进行目标的数目统计,因此,需要预先将带有人为标定目标中心位置的遥感图片转变成真值密度图,再进行训练。在对整个网络的训练中,为了评估经网路所估计到的密度图与真值密度图之间的差异,需要对目标函数(损失函数)进行优化。最后,在测试阶段,为了评估本发明的方法的有效性,采用经典的评价指标对所提方法进行评估。具体地,

[0057] 关于真值密度图的生成:假设像素位置为 $x_n$ (目标中心坐标)的一个目标,能够用一个脉冲函数 $\delta(x-x_n)$ 表示,对于一幅含有N个目标的图像,则可以表示为:

$$H(x) = \sum_{n=1}^{N} \delta(x - x_n)$$
,为了产生密度图F,将 $H(x)$ 与高斯核进行卷积,即 $F(x) = H(x) * G_{\sigma_n}(x)$ ,

其中,H(x) 为表示含有N个目标的图像的函数,F(x) 为真值密度图函数, $G_{\sigma_n}(x)$ 为方差为 $\sigma_n$ 的高斯函数,n表示第n个目标, $\sigma_n$ 表示标准差,设定固定核 $\sigma_n=15$ 。

[0058] 关于损失函数:采用欧式距离函数作为评估预测密度图和真值密度图之间的差异的损失函数,损失函数 $L(\Theta) = \frac{1}{2B} \sum_{b=1}^{B} \left\| F(X_b; \Theta) - F_b^{GT} \right\|_2^2$ ,其中,B表示批量大小,X<sub>b</sub>表示输入图

片,b表示第b幅图像, $\Theta$ 表示训练的参数, $F(X_b;\Theta)$ 和 $F_b^{GT}$ 分别表示估计的密度图和对应的 真值密度图。

[0059] 为了评估本发明方法的有效性,采用两个评价指标来进行评估:平均绝对误差 (Mean Absolute Error, MAE) 和均方误差(Mean Squared Error, MSE), MAE评价模型的准确 性,MSE用来评测模型的鲁棒性。两个指标的定义为:

$$MAE = \frac{1}{T} \sum_{t=1}^{T} \left| C_t^{pred} - C_t^{GT} \right|$$
 ,  $MSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left| C_t^{pred} - C_t^{GT} \right|^2}$  ,其中,T表示测试样本的数目,t为

第t幅图像, $C_t^{pred}$ 和 $C_t^{GT}$ 分别表示估计的目标数和真实的目标数。为了方便理解本发明的上 述技术方案,以下通过具体实施例对本发明的上述技术方案进行详细说明。

[0060] 实施例1

[0061] 在包含3057幅图片,由建筑物、小车、大型货车、船只4类目标组成的数据集上验证 本发明提出的方法,数据集的具体数据统计可见表1。

[0062] 表1 验证本发明所使用的数据集信息统计

数据集 图像数目 训练/测试 分辨率 数目统计 平均 总数 最小 最大 建筑物 1205/1263 76, 215 2468  $512 \times 512$ 15 30.88 142 [0063] 小车 280 222/58  $2473 \times 2339$ 148,838 17 531. 56 8531 大型货车 172 108/64  $1552 \times 1573$ 16, 594 12 96.48 1336 船只 137 97/40 $2558 \times 2668$ 44,892 50 327.68 1661

[0064] 如图6(a)-图6(c),图7(a)-图7(c),图8(a)-图8(c),图9(a)-图9(c),本发明的模 型是端到端进行训练的,网络的前10层是在VGG16网络结构进行精调的,其他卷积层的参数 采用0.01的标准偏差的高斯初始化。在训练期间,采用随机梯度下降(stochastic gradient descent, SGD), 学习率设定为1e-5。对于建筑物数据集,采用批量大小为32,经过 400个周期直至训练收敛:对于其他三类,即船只、小车和大型车辆数据集的,采用批量大小 为1,同样历经400个周期进行训练。

[0065] 为了扩增训练集,在图片的不同位置裁剪9块图像块,每块的尺寸为原始图片的1/ 4,前4块是不重叠的图像块,后五块是随机位置进行裁剪的,然后将这些块进行水平翻转。 因为船只、小车和大型车辆数据集的图片的分辨率比其他常规数据集的图片大,很容易导 致显卡内存不足。因此,在数据增强之前,先将这些图片的尺寸固定为1024×768。模型是用 pytorch编写,在NVIDIA GTX 2080Ti GPU上进行实验。

[0066] 为了验证模型每个模块的有效性,在建筑物数据集上进行消融实验。实验过程包 括基准实验和连续在此基础上添加三个模块:

● 基准实验:采用CSRNet作为基准方法(前端网络以VGG16网络结构为主干网络, 后端网络采用6层膨胀因子为2的卷积层);

[8600] ● 基准+注意力模块:在基准方法的基础上,添加连接通道注意力机制和空间注 意力机制的模块;

● 基准+注意力模块+尺度金字塔模块:在前面基础上添加尺度金字塔模块; [0069]

[0070] ● 基准+注意力模块+尺度金字塔模块+可变形卷积模块:本发明提出的方法。

[0071] 消融实验的结果如表2所示,从表中可知,本发明的网络中的每个模块对性能的提升都有一定的贡献。具体地,原始的基准方法在数据集上表现并不是很理想,添加上注意力模块后,采集特征图全局和局部的依赖信息,使得性能有了一定的提升;再添加上尺度金字塔模块后,性能得到进一步的提升;最后,融合可变形卷积后,本发明提出的模型在数据集上表现出最优的性能。

[0072] 表2 在建筑物数据集上的消融实验

	方法        建筑物数据组					
		MAE	MSE			
[0073]	基准方法	8.00	11.78			
	基准方法+注意力模块	7.92	11.67			
	基准方法+注意力模块+尺度金字塔模块	7.85	11.58			
	本发明的方法	7.59	10.66			

[0074] 表3为本发明的方法与其他方法的结果对比。这些方法包括:MCNN、CMTL、CSRNet、SFCN、SANet、SPN、SCAR。从表中发现,本发明的方法在构建的遥感目标计数数据上表现出最优的结果,同时也说明本发明的方法具有良好的泛化能力。

[0075] 表3 本发明的方法与其他方法的对比

[0076]	数据集方法	建筑物		小车		大型车辆		船只	
		MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
[0077]	MCNN	13. 65	16. 56	488. 65	1317. 44	36. 56	55. 55	263. 91	412. 30
	CMTL	12. 78	15. 99	490. 53	1321. 11	61. 02	78. 25	251. 17	403. 07
	CSRNet	8.00	11. 78	443. 72	1252. 22	34. 10	46. 42	240. 01	394. 81
	SANet	29. 01	32. 96	497. 22	1276. 66	62. 78	79. 65	302. 37	436. 91
	SFCN	8. 94	12. 87	440. 70	1248. 27	33. 93	49. 74	240. 16	394. 81
	SPN	7. 74	11. 48	445. 16	1252. 92	36. 21	50. 65	241. 43	392. 88
	SCAR	26. 90	31. 35	497. 22	1276. 65	62. 78	79. 64	302. 37	436. 92
	ASPDNet(Ours)	7. 59	10.66	433. 23	1238. 61	18. 76	31.06	193. 83	318. 95

[0078] 在本发明中,术语"第一"、"第二"、"第三"、"第四"仅用于描述目的,而不能理解为指示或暗示相对重要性。术语"多个"指两个或两个以上,除非另有明确的限定。

[0079] 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,对于本领域的技术人员来说,本发明可以有各种更改和变化。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。

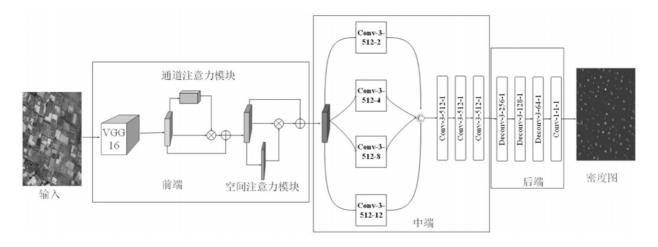


图1

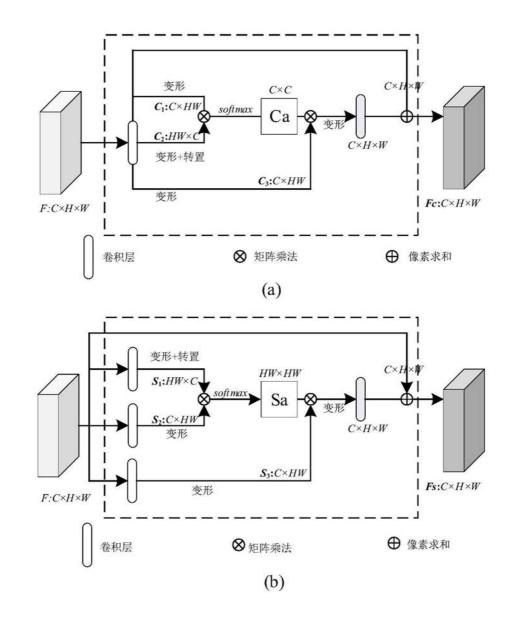
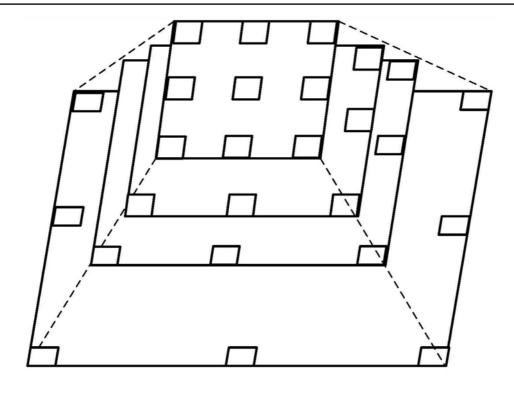


图2



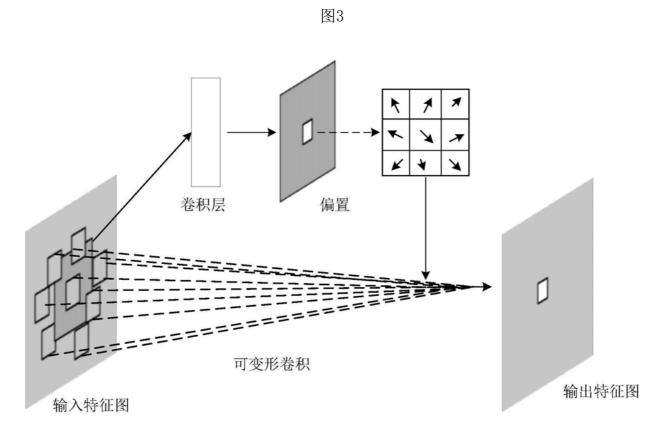


图4

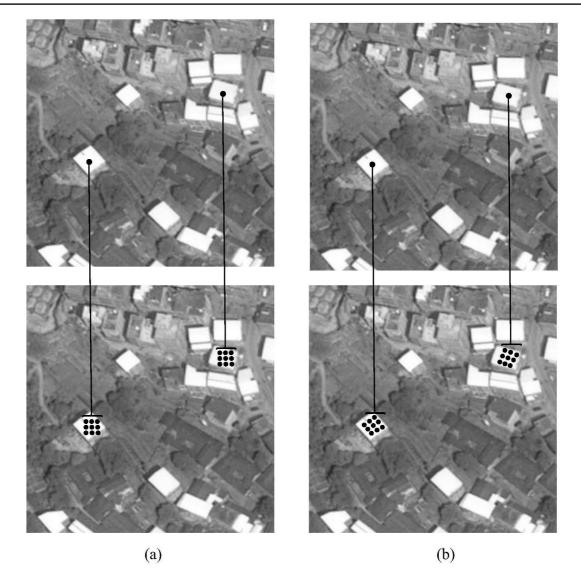


图5

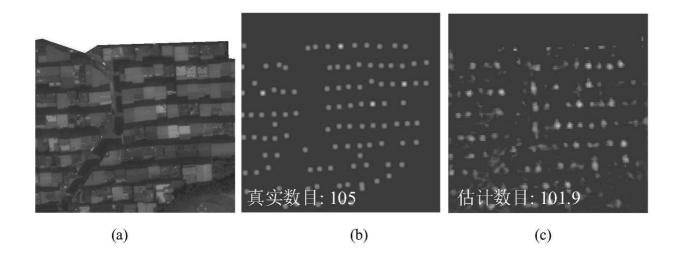


图6

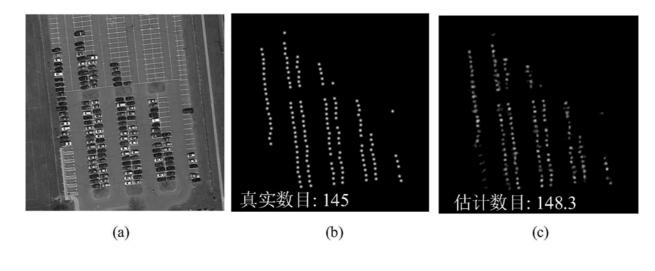


图7

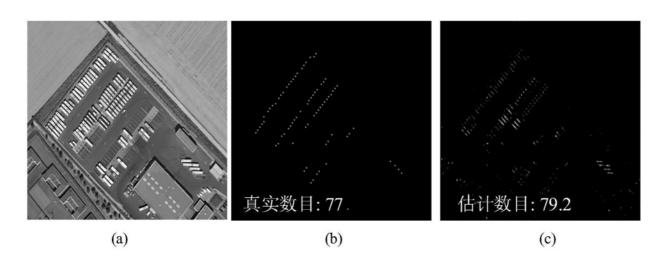


图8

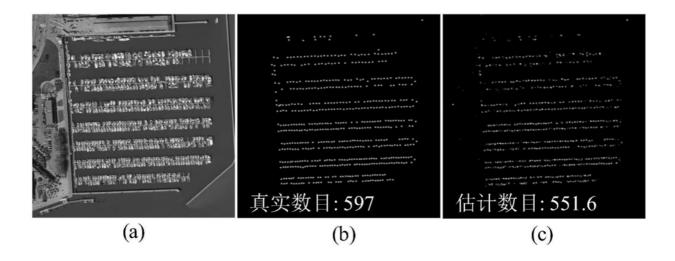


图9