

一种面向航空图像的自适应目标计数模型

魏畅, 关侏红, 张毅超, 李文根

引用本文

魏畅, 关侏红, 张毅超, 李文根. 一种面向航空图像的自适应目标计数模型[J]. 计算机科学, 2023, 50(8): 93-98.

WEI Chang, GUAN Jihong, ZHANG Yichao, LI Wengen. [Adaptive Object Counting Model for Aerial Imagery](#) [J]. Computer Science, 2023, 50(8): 93-98.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[双编码半监督异常检测模型](#)

Dually Encoded Semi-supervised Anomaly Detection

计算机科学, 2023, 50(7): 53-59. <https://doi.org/10.11896/jsjcx.220900027>

[基于暗通道先验的单幅图像去雾新算法](#)

Novel Algorithm of Single Image Dehazing Based on Dark Channel Prior

计算机科学, 2021, 48(7): 219-224. <https://doi.org/10.11896/jsjcx.200700160>

[基于Bloom Filter的安全P2P共享模型](#)

Security P2P Sharing Model Based on Bloom Filter

计算机科学, 2010, 37(5): 72-76.

[范例推理结构相似性研究](#)

计算机科学, 2004, 31(6): 153-155.

[基于Haar小波变换高频特征的图像质量评价算法](#)

Image Quality Assessment Algorithm Based on High Frequency Band of Haar Wavelet Transform

计算机科学, 2009, 36(10): 262-264.

一种面向航空图像的自适应目标计数模型

魏 畅 关侏红 张毅超 李文根

同济大学电子与信息工程学院 上海 201800

(593355341@qq.com)

摘 要 目标计数旨在获取给定图像中包含的车辆、建筑物、人物等特定种类目标的数量,对城市规划、应急响应、国家安全等具有重要意义。当前目标计数任务主要依赖于低空摄像头所拍摄的图像,存在目标易被遮挡和计数空间范围小等突出问题。高清航空遥感图像的广泛使用使大范围目标计数成为可能。然而,面向航空图像的目标计数任务存在目标尺度差异大、分布密集、方向不确定等挑战,现有基于低空图像的目标检测计数模型和回归计数模型均无法适用于航空图像的目标计数。针对该问题,提出了一种面向航空图像的自适应目标计数模型。首先,利用几何自适应高斯卷积方法解决目标尺度变化问题;然后,利用基于结构相似性的图片损失判断方法解决目标密集区域计数稳定性较差的问题。实验结果表明,所提模型相较于基准模型取得了更好的目标计数精度。

关键词: 目标计数;航空图像;回归计数;高斯卷积;结构相似性

中图分类号 TP391

Adaptive Object Counting Model for Aerial Imagery

WEI Chang, GUAN Jihong, ZHANG Yichao and LI Wengen

School of Electronic and Information Engineering, Tongji University, Shanghai 201800, China

Abstract Object counting aims to obtain the number of specific types of objects such as vehicles, buildings, people contained in a given image, which is of great significance to urban planning, emergency response, national security, etc. The current object counting task mainly relies on the images taken by low-altitude cameras, and there are obvious problems such as the object being easily occluded and the small counting space range. Widespread use of high-definition aerial remote sensing imagery makes it possible to count objects in large areas. However, the object counting task for aerial images has challenges such as large differences in object scales, dense distribution, and uncertain orientation. Existing object detection counting models and regression counting models based on low-altitude images are not suitable for object counting in aerial images. To solve this problem, this paper proposes an adaptive object counting model for aerial images. Firstly, the geometric adaptive Gaussian convolution method is used to solve the problem of object scale variation. Then, the image loss judgment method based on structural similarity is used to solve the problem of poor counting stability of object dense regions. Experimental analysis shows that the proposed model can achieve better object count accuracy than the benchmark model.

Keywords Object counting, Aerial imagery, Regression counting, Gaussian convolution, Structural similarity

1 引言

随着城市化水平不断提高,城市中车辆、船舶、建筑物等地理物体的数量愈发增多,目标之间更加拥挤,城市的地面结构也更加复杂。要对城市资源进行合理的规划与调整,就需要对这些地理结构有一个系统全面的掌握。进行目标计数任务,对这些目标的数量进行准确估计是实现这一目标的关键。例如,车辆计数可以辅助城市道路规划,建筑物计数可以辅助城市空间设计。

具体而言,目标计数任务是指估计一定区域内指定单个

类别或多个类别目标的数量,计数目标包括人群、车辆、船舶与细胞等。与目标检测任务不同,计数任务减少了对目标位置的关注,注重最后计数结果的准确度,是一项重要且实用的计算机视觉任务。其在人群计数^[1]、细胞观察^[2]、环境调查^[3]等领域都有着广泛的应用。

然而,现有目标计数方法主要依赖于设置在路口或低空的摄像头所拍摄的图像。这些图像存在诸多问题,首先,由于摄像头角度的限制,目标很容易被障碍物遮挡;其次,由于摄像头方向的单一性和拍摄距离的限制,此类方法只能观测局部信息,计数空间范围小,不能进行大范围评估。

到稿日期:2022-06-18 返修日期:2022-11-14

基金项目:国家自然科学基金联合基金(U1936205)

This work was supported by the Joint Funds of the National Natural Science Foundation of China(U1936205).

通信作者:关侏红(jhguan@tongji.edu.cn)

当前,遥感卫星的侦查范围遍布全球,可以及时获取各地的地面信息。随着光学遥感检测技术的不断发展,航空遥感图像分辨率也逐步提高,这使得利用高质量航空图像进行地面目标计数成为可能。光学遥感的航空图像能覆盖更广阔的空间范围,进行大范围地面监测,有助于获取全局信息。此外,航空图像的垂直视角特性减小了目标被遮挡的可能性,可以提高目标的识别率。如图1所示,遥感卫星拍摄的航空图像可以清晰显示大范围的车辆信息。与传统摄像图像相比,航空图像存在目标尺度差异大、分布密集、方向不确定、背景杂乱等特点。



图1 包含大量车辆的航空图像

Fig. 1 Aerial image with a large number of cars

现有目标计数方法主要有两大类:基于检测的方法与基于回归的方法。

早期基于检测的模型^[4-7]以模版匹配技术为基础,其基本结构包括区域选择器、特征提取器和分类器3个部分。首先,模型使用区域选择器的“选择性搜索”方法,查找图像中可能表示目标的区域,再利用特征提取器提取这些区域的特征,最后利用分类器与现有模版进行匹配,完成目标分类。这些模型依赖于人工设计的特征,特征泛化能力较弱,缺乏提取高级语义特征的能力。基于深度学习的R-CNN^[8-10]系列检测模型能够学习到更深层次的图像特征,优化了模型的整体框架结构,能有效提高运算效率和目标计数结果的精度。后续的工作中,YOLO^[11],SSD^[12]等去掉了生成候选框的步骤,直接对原图回归,大大加快了检测速度。最近的工作中,基于对比

学习的半监督方法^[13-14]和基于自注意力的方法^[15-16]成为研究的热点,使检测的精度与效率得到进一步提升。然而此类方法依然是以检测为基础,需要给目标划定清晰的边界,在面对分布密集的目标时,容易产生检测目标重叠的问题,导致检测性能较差。

基于回归的模型可以进一步划分成回归计数法和密度图估计法。回归计数法^[17]输入目标图片进入深度网络,通过端到端学习直接预测目标数量,适用于目标密度较为稀疏的图片。而在密度图估计法^[18]中,模型输出预测密度图,通过对密度图进行逐位积分求和得到最终的目标数量。最新研究中对半监督和自注意力方法^[19-20]的使用,也让基于回归的目标计数模型达到了较高的准确率。但通用的目标计数方法缺少对遥感目标的方向处理模块,影响了计数精度。

目前,针对航空图像的目标计数的研究中,ASPDNet^[21]使用注意力机制与可变形卷积解决背景杂乱以及航空图像中目标方向不确定性的问题,TDME^[22]模型利用Swin Transformer^[23]结构,生成了较大的感受野,有效提取了遥感图像的特征。但是,这类方法仍然没有解决目标尺度差异大以及分布密集的问题。

针对上述问题,本文提出了一种面向航空图像的自适应目标计数模型(Adaptive Object Counting Model for Aerial Imagery, AOCM);并提出了一种几何自适应高斯卷积方法,用于解决目标尺度差异问题;提出了一种结构损失性损失函数,解决了密集区域计数稳定性的问题。本文提出的方法有效提升了航空图像目标计数的准确度。

2 面向航空图像的自适应目标计数模型

图2给出了AOCM模型的基本结构。该模型主要包含几何自适应高斯卷积模块、回归计数模块和结构相似性损失函数模块3个模块。首先,几何自适应高斯卷积模块生成真实密度图,并通过回归计数模块生成预测密度图,再通过结构相似性损失函数模块对真实密度图与生成密度图的相似度进行评估,进而优化网络参数,最终完成目标计数任务。

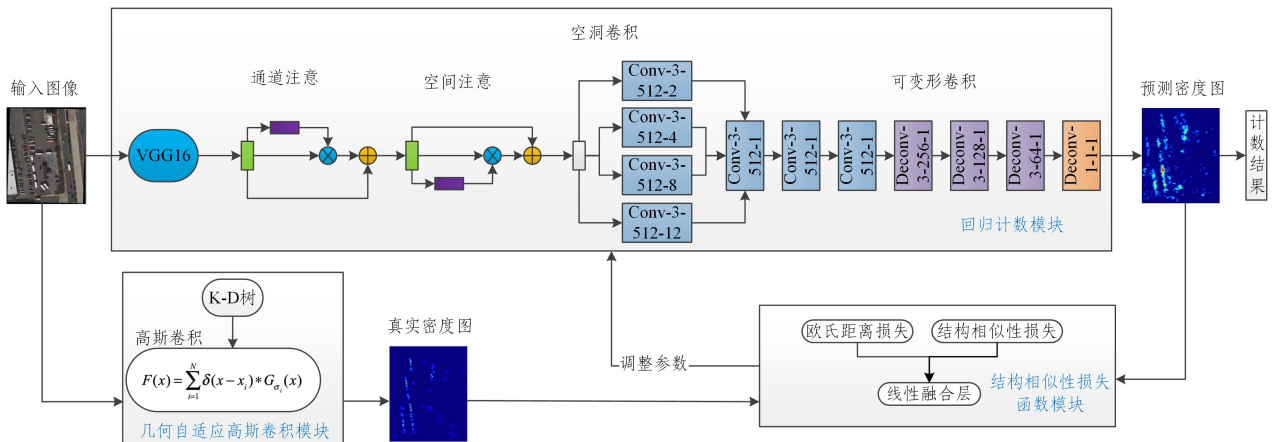


图2 AOCM模型网络结构

Fig. 2 Structure of AOCM model

AOCM模型的回归计数模块参考了Gao等提出的ASPDNet^[21]模型。如图2所示,模块分为3个子模块:前端模块利用通道注意力机制与空间注意力机制捕获目标的特征,

解决航空图像中拥挤的小目标难以与背景区分的问题;中间模块利用空洞卷积应对目标比例的变化,解决航空图像中同种目标比例变化大、大小差异明显的问题;后端模块利用

可变形卷积处理方向信息,解决垂直视角拍摄所导致的航空图像方向不确定的问题。

下面将以车辆计数任务为例,分别介绍几何自适应高斯卷积与结构相似性损失函数的功能、结构和技术细节。

2.1 几何自适应高斯卷积模块

航空图像数据集图片的目标标注一般使用中心点标注的方式,在图片中仅标注出每个待计数目标中心点的位置。因此,在进行学习时,无法通过单个孤立的点学习目标特征,故对于每个标注点,一般采用高斯卷积的方式来生成密度图。高斯卷积采用了高斯模糊的原理。

$$F(x) = \sum_{i=1}^N \delta(x - x_i) * G_{\sigma_i}(x) \quad (1)$$

其中, δ 代表delta函数,该函数在除了0以外的点取值都等于0,而其在整个定义域上的积分等于1; G_{σ_i} 代表高斯核函数; σ_i 代表高斯核标准差; x_i 代表目标实例, N 代表目标数量。

高斯模糊令每一个像素取周边像素的加权平均值,把图像中单个点标注的影响扩张到整个密度图。高斯模糊生成的密度图效果与高斯核选取的标准差相关,高斯核的标准差越大,图片表现出的模糊效果就越强烈。

现有方法在对图像进行处理时,采用固定高斯核法,对所有目标点使用统一的高斯核标准差 σ 。这种方法忽略了目标的尺度差异,计算速度较快,面对尺度差异不大的目标时,可以取得较好的效果。但是,在检测如车辆这类密集的小物体时,因为不同图像中目标尺度差异过大,容易产生高斯核设置过大而导致密度图中的目标重叠,或是高斯核设置过小导致目标被忽略的问题,从而影响检测的结果。

AOCM模型利用几何自适应高斯卷积模块来解决这种尺度差异问题,通过待检测物体的大小来确定高斯核的大小,让模糊效果与目标大小相适应。但是,由于数据集图片中的检测目标仅标记了中心点,因此目标的大小无法直接确定。考虑到航空图像中目标的排列较为密集,本文采用目标间距离估计目标大小来间接确定高斯核标准差 σ_i 。

为实现高效的目标间距离估计,本文利用K-D树划分图像。K-D树结构如图3所示。

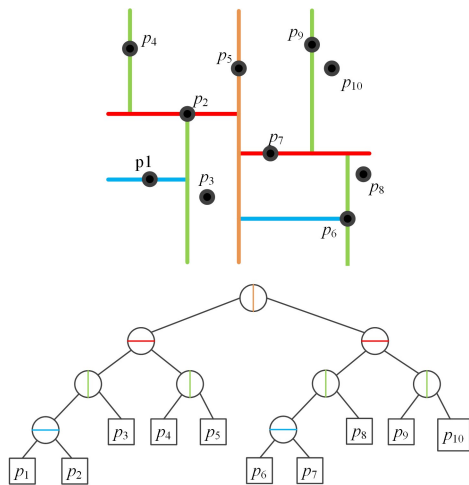


图3 目标空间划分和K-D树结构

Fig. 3 Spatial partition of objects and the structure of K-D tree

首先,在图像中选择中心目标点为端点,从端点依次用

垂直线和水平线来分割空间中的目标,平面一侧的点归为左子树,另一侧的点归为右子树,将平面上的无序点划分出一定的层次结构。

然后,对所有目标搜寻与其距离最近的3个目标,计算与这3个目标的平均距离,用该距离来估计目标大小,从而对图像中的每一个计数目标自适应地确定高斯核标准差的大小。

$$\sigma_i = \lambda_1 (L_{i1} + L_{i2} + L_{i3}) \quad (2)$$

其中, σ_i 代表对目标*i*生成的自适应高斯核标准差, λ_1 代表线性参数, L_{i1}, L_{i2}, L_{i3} 分别代表目标*i*与最邻近3个目标点之间的距离。

针对密集区域的对象,几何自适应法可以随着目标间距离调整核大小,效果优于固定高斯核法。但是,在面对稀疏目标时,由于目标间距离较大,根据距离估计目标大小的方法不够准确,容易导致高斯核过大,造成密度图模糊。因此,本文设置了一定的阈值 σ ,如果对于物体大小的估计超过了这一阈值,则采用原有固定高斯核,这样可以在一定程度上照顾到远处的孤立物体,以免其在生成密度图的过程中被忽视。

相比固定高斯核,几何自适应高斯核在处理较为密集且尺度差异较大的物体时,可以通过目标大小自适应地调整高斯核的大小,有效地应对尺度差异,在密集的区域降低模糊性,产生更加清晰的边界,有利于后续的目标数量估计。

几何自适应高斯卷积算法如算法1所示。

算法1 几何自适应高斯卷积算法

输入:输入图像I,标注点集 $P = \{(x_1, y_1), (x_2, y_2), \dots, (x_p, y_p)\}$, 阈值 σ ,线性参数 λ_1

输出:输出密度图像F

1. $(h, w) \leftarrow \text{Size}(I)$ // 获取尺寸
2. $F \leftarrow \text{New}(h, w)$ // 新建图
3. $i = 0, j = 0, \text{Tree} = []$
4. WHILE $i < h$;
5. WHILE $j < w$;
6. IF $(i, j) \text{ IN } P$;
7. $\text{Tree} \leftarrow \text{Tree} + (i, j)$
8. $\text{Near} \leftarrow \text{KD}(\text{Tree})$ // K-D树获取临近点
9. FOR $p \text{ IN } \text{Tree}$ DO;
10. $\text{distance} \leftarrow \text{Euclidean}(p, \text{Near})$
11. $\sigma_p \leftarrow \lambda_1 * \text{Sum}(\text{distance})$
12. IF $\sigma_p > \sigma$;
13. $\sigma_p \leftarrow \sigma$
14. $F \leftarrow F + \text{Gaussian}(p, \sigma_p)$ // 高斯卷积
15. 结束输出密度图F

2.2 结构相似性损失函数

针对航空图像的目标计数中,受密集区域影响计数稳定性较差的问题,AOCM模型引入基于结构相似性度量的损失函数。

$$Loss = \frac{1}{N} \sum_{i=1}^N (L_i + \lambda_2 (1 - L_i^{SSIM})) \quad (3)$$

其中, L 代表预测密度图像与真实密度图像之间的欧氏距离损失,SSIM是结构相似性损失, λ_2 是线性参数, N 是图像数量。

欧氏距离损失*L*的计算公式如下:

$$L_i = \|F(X_i) - F_i^{GT}\|_2^2 \quad (4)$$

其中, $F(X_i)$ 代表预测密度图像, F_i^{GT} 代表真实密度图像。通过对每个点之间的欧氏距离取平均来评估图像之间的差异。对于计数任务来说, 该损失项与最终计数的目标相重合, 可以得到较好的准确率。该损失项不关注图像的结构化信息, 对于整体图像结构信息的相似程度不够敏感, 个别像素点的极端情况对整体计数效果影响很大。

为克服欧氏距离损失的不足, AOCM 模型引入结构相似性损失 SSIM^[24] 来度量密度图的质量。SSIM 通过均值、方差、协方差 3 个局部统计量来计算两张图像之间的相似度。

$$L_i^{SSIM} = (2\mu_e\mu_g + c_1)(2\sigma_{eg} + c_2)(\mu_e^2 + \mu_g^2 + c_1)(\sigma_e^2 + \sigma_g^2 + c_2) \quad (5)$$

其中, μ_e, μ_g 分别代表两张图像的均值; σ_e 和 σ_g 分别代表两张图像的方差; σ_{eg} 代表两张图像间的协方差; c_1, c_2 为平滑系数, 避免分母为 0。结构相似性损失能够更好地关注图像间的局部差异, 从而更好地生成密度图。

AOCM 模型将欧氏距离损失和结构相似性损失线性结合, 其在使用欧几里得距离保证计数结果准确性的同时, 也用 SSIM 来考虑图像结构化信息, 提高生成图像的稳定性。

3 实验评估

3.1 评价指标

本文使用平均绝对误差 (Mean Absolute Error, MAE) 和均方根误差 (Mean Squared Error, RMSE) 两个指标来评价目标计数方法的性能。两个评价指标的计算公式如下:

$$MAE = \frac{1}{N} \sum_{i=1}^N |C_{I_i}^{pred} - C_{I_i}^{gt}| \quad (6)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N |C_{I_i}^{pred} - C_{I_i}^{gt}|^2} \quad (7)$$

其中, N 为测试图片的数量, $C_{I_i}^{pred}$ 和 $C_{I_i}^{gt}$ 分别代表图片目标数量的预测结果和真实值。严格来说, MAE 反映了预测的准确性, RMSE 反映了预测的稳定性和鲁棒性。

3.2 实验配置

本文以车辆计数任务为例验证所提方法的有效性。目前, 车辆航空图像数据集主要用于目标检测与分类任务, 常见数据集有 COWC^[25] 数据集、CARPK^[26] 数据集等。这些数据集都存在背景杂乱、实例种类多、单张图片车辆数量少等特点。本文针对车辆计数的特点, 从 DOTA^[27] 数据集中选取了一定数量的图片进行实验。DOTA 数据集是从中国资源卫星数据和应用中心提供的卫星图像以及航拍图像中收集的图片, 包含 2 806 张航空图像, 尺寸大约为 $4\,000 \times 4\,000$, 包含 15 个类别共计 188 282 个实例, 是一个用于航空目标检测的大规模数据集。本文从 DOTA 数据集中选取了 90 张适合车辆计数任务的图片, 并随机选取 65 张图片作为训练集, 25 张作为测试集。尽管数据集包含的图片数量较少, 但每张图片均包含数百辆汽车, 计数目标分布密集, 能够较好地检验模型的准确程度。此外, 数据集涵盖了停车场、城镇、村庄等各种场景, 能充分评估模型在不同背景条件下的性能。

为适合回归计数的流程, 本文采取中心点图片标注方式。中心点标注基于 DOTA 数据集给定的目标注释的 4 个

点坐标, 连接坐标形成一个包围待检测目标的四边形, 用该四边形的中心点对物体进行标记。该标注方式仅考虑目标的位置特点, 忽略目标的大小与方向, 可以加快学习速度, 以适应后续任务中密度图的生成。

表 1 列出了实验环境的各项参数。在模型训练时, 根据实验分析, 前文公式中的线性参数 λ_1 和 λ_2 分别设置为 0.1 和 1.0。

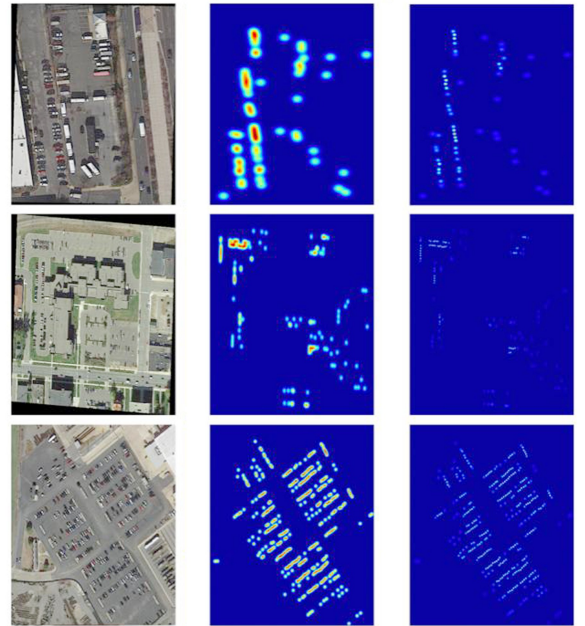
表 1 实验环境

Table 1 Experimental environment

| | |
|-----------|--------------------------------------|
| 操作系统 | CentOS Linux release 7.8.2003 (Core) |
| 显卡型号 | GeForce RTX 3090 (×2) |
| 实际显存大小/MB | 24 268 |

3.3 几何自适应高斯卷积结果

固定高斯核和几何自适应高斯核生成密度图的具体对比结果如图 4 所示。可以看出, 在目标较为密集的区域, 固定高斯核产生了目标杂糅现象, 难以分辨目标具体的位置。自适应高斯核在密集区域自适应调整高斯核的大小, 对单个目标产生了更为清晰的边界, 有利于后续的目标数量估计。



(a) 原遥感图像 (b) 固定高斯核 (c) 几何自适应高斯核

图 4 固定高斯核和几何自适应高斯核的密度图生成结果

Fig. 4 Results of density map generation of fixed Gaussian kernel and geometrically adaptive Gaussian kernel

3.4 方法对比实验结果

本文选取的进行对比的回归计数模型为 Gao 等提出的 ASPDNet^[21] 网络。此外, 还对传统的目标检测网络 SSD^[12] 进行了适应性改进, 将其作为参照来比较回归方法和检测方法在目标计数时的效果差异。

表 2 列出了各类模型的实验结果。从表中可以看出, SSD 检测方法由于检测框大小固定, 没有额外的小目标关注模块, 对小目标的识别性较差。并且, 由于使用固定的水平检测框, SSD 方法无法应对航空图像中的目标方向不确定性, 因此该方法的 MAE 与 RMSE 显著高于另外两种方法。适用于大型建筑计数的回归计数方法 ASPDNet 利用注意力模块

与可变形卷积模块解决了小目标识别问题与目标方向不确定性问题,取得了不错的计数效果。AOCM 提高了应对目标尺度变化的能力与计数稳定性,在车辆计数数据集上的效果得到了进一步的提升,在 3 种方法中取得了最优的实验结果。

表 2 对比实验结果

Table 2 Experimental results comparison

| 算法模型 | MAE | RMSE |
|---------|---------------|---------------|
| ASPDNet | 70.472 | 96.446 |
| SSD | 125.192 | 157.639 |
| AOCM | 68.199 | 91.717 |

图 5 对预测效果进行了可视化,给出了不同图片的车辆密度图生成结果。从预测的直观效果来看,AOCM 方法生成的密度图结构分层更为清晰,与原图相比,相似度也更高。这是因为 AOCM 引入的几何自适应高斯卷积模块与结构相似性损失函数模块让模型更关注图像整体的结构,提高了所生成密度图的整体稳定性。

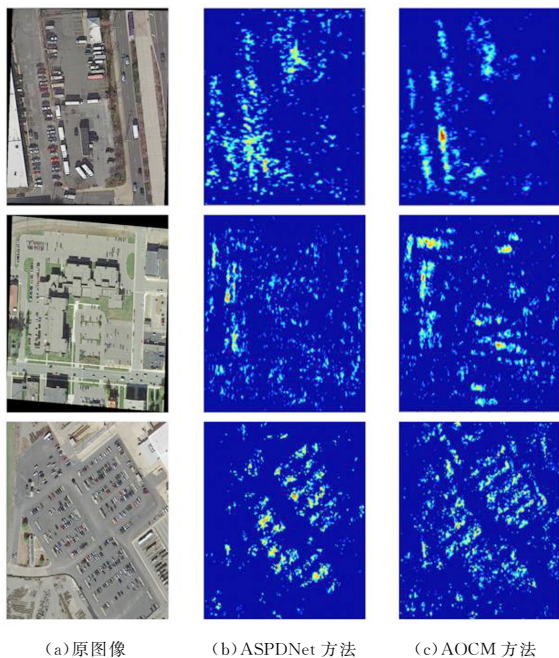


图 5 预测效果可视化

Fig. 5 Visualization of predicted results

3.5 模块有效性分析实验结果

为了进一步验证几何自适应高斯卷积和结构相似性损失函数在网络中的有效性,本文设计了如下 3 组对比实验:

- (1) AOCM 模型去掉结构相似性损失函数模块;
- (2) AOCM 模型去掉几何自适应高斯卷积模块;
- (3) 完整的 AOCM 模型。

表 3 列出了 3 组实验的结果。对比实验结果可以看出,几何自适应高斯卷积模块对整体效果有一定的提升,因为自适应的卷积核提高了生成密度图的清晰度,有利于模型的学习过程。结构相似性损失函数的优化对于整体的效果也有明显提升,尤其是在 RMSE 上的提升最明显。这是因为加入结构相似性误差在一定程度上缓解了单纯欧氏距离损失中容易被个别的极端像素点影响的问题,提高了模型的稳定性。

表 3 模块有效性分析实验结果

Table 3 Effectiveness analysis of modules

| 实验组号 | MAE | RMSE |
|------|---------------|---------------|
| 1 | 69.356 | 96.213 |
| 2 | 68.565 | 92.890 |
| 3 | 68.199 | 91.717 |

结束语 为解决航空图像中目标计数存在的挑战,本文提出了一种面向航空图像的自适应目标计数模型 AOCM,利用几何自适应高斯卷积方法解决了目标尺度差异大而导致计数准确性较差的问题,利用基于结构相似性的损失函数解决了目标密集区域计数稳定性较差的问题。通过实验分析可知,AOCM 模型在基于航空图像的车辆计数问题上取得了较好的计数效果。但由于航空图像中目标分布不均,密集区域相对于稀疏区域对整体计数结果的影响更大;而基于高斯卷积的模型受稀疏区域影响较大,在面临分布差异较大的图像中计数效果不佳。在后续研究中可以尝试在模型中引入空间聚类与注意力机制,增加对密集区域的关注程度。

参考文献

- [1] ZHANG Y,ZHOU D,CHEN S,et al. Single-image crowd counting via multi-column convolutional neural network[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:589-597.
- [2] WANG Y,ZOU Y. Fast visual object counting via example-based density estimation[C]// 2016 IEEE International Conference on Image Processing(ICIP). IEEE,2016:3653-3657.
- [3] FRENCH G,FISHER M,MACKIEWICZ M,et al. Convolutional neural networks for counting fish in fisheries surveillance video[C]// Workshop on Machine Vision of Animals and their Behaviour. BMVA Press,2015:7. 1-7. 10.
- [4] DOLLAR P,WOJEK C,SCHIELE B,et al. Pedestrian detection: a benchmark[C]// 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE,2009:304-311.
- [5] UIJLINGS J R,VAN D S K E A,GEVERS T,et al. Selective search for object recognition[J]. International Journal of Computer Vision,2013,104(2):154-171.
- [6] LAZEBNIK S,SCHMID C,PONCE J. Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories[C]// Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Paris:IEEE,2006:2169-2178.
- [7] YU Y,ZHANG J,HUANG Y,et al. Object detection by context and boosted HOG-LBP[C]// ECCV Workshop on PASCAL VOC. 2010.
- [8] GIRSHICK R. Fast R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. Santiago:IEEE,2015:1440-1448.
- [9] REN S,HE K,GIRSHICK R,et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(6):1137-1149.
- [10] HE K,GKIOXARI G,DOLLAR P,et al. Mask R-CNN[C]// Proceedings of the IEEE Conference on Computer Vision and Pat-

- tern Recognition. Venice;IEEE,2017;2961-2969.
- [11] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu;IEEE,2017;7263-7271.
- [12] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multi Box Detector[C]//European Conference on Computer Vision. Cham;Springer,2016;21-37.
- [13] XIE E, DING J, WANG W, et al. Detco: Unsupervised contrastive learning for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021;8392-8401.
- [14] XU M, ZHANG Z, HU H, et al. End-to-end semi-supervised object detection with soft teacher[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021;3060-3069.
- [15] DAI X, CHEN Y, YANG J, et al. Dynamic detr: End-to-end object detection with dynamic attention[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021;2988-2997.
- [16] DAI X, CHEN Y, XIAO B, et al. Dynamic head: Unifying object detection heads with attentions[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021;7373-7382.
- [17] WANG C, ZHANG H, YANG L, et al. Deep people counting in extremely dense crowds[C]//Proceedings of the 23rd ACM International Conference on Multimedia. New York;ACM,2015;1299-1302.
- [18] ZHANG C, LI H S, WANG X G, et al. Cross-scene crowd counting via deep convolutional neural networks [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ;IEEE,2015;833-841.
- [19] LIANG D, CHEN X, XU W, et al. TransCrowd: weakly-supervised crowd counting with transformers[J]. Science China Information Sciences,2022,65(6):1-14.
- [20] SUN G, LIU Y, PROBST T, et al. Boosting crowd counting with transformers[J]. arXiv:2105.10926,2021.
- [21] GAO G, LIU Q, WANG Y. Counting dense objects in remote sensing images [C] // ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP). IEEE,2020;4137-4141.
- [22] DING G, CUI M, YANG D, et al. Object Counting for Remote Sensing Images via Adaptive Density Map-Assisted Learning [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022,60;1-11.
- [23] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021;10012-10022.
- [24] CAO X, WANG Z, ZHAO Y, et al. Scale aggregation network for accurate and efficient crowd counting[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018;734-750.
- [25] MUNDHENK T N, KONJEVOD G, SAKLA W A, et al. A large contextual dataset for classification, detection and counting of cars with deep learning[C]//European Conference on Computer Vision. Cham;Springer,2016;785-800.
- [26] HSIEH M R, LIN Y L, HSU W H. Drone-based object counting by spatially regularized regional proposal network [C] // Proceedings of the IEEE International Conference on Computer Vision. 2017;4145-4153.
- [27] XIA G S, BAI X, DING J, et al. DOTA: A large-scale dataset for object detection in aerial images[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018;3974-3983.



WEI Chang, born in 1999, master, is a student member of China Computer Federation. His main research interests include target counting, spatiotemporal data management and analysis.



GUAN Jihong, professor, doctoral supervisor, is a member of China Computer Federation. Her main research interests include artificial intelligence, spatiotemporal data management and analysis, network modeling, game and collaboration, application-oriented big data analysis and application services.

(责任编辑:何杨)