

基于 GN 分裂的小目标检测区域推荐搜索算法

赵沛然^{**}, 吴新元, 汤新雨, 沈晓海, 许海燕, 李敏, 张学武^{*}

河海大学物联网工程学院, 江苏 常州 213022

摘要 区域推荐搜索是机器视觉研究热点之一, 针对传统目标检测使用穷举式搜索效率低下的问题, 通过优化搜索的准确率可提高检测效率。引入复杂网络中用于社区发现的 Girvan-Newman(GN)分裂算法, 结合小目标区域特征, 提出一种基于图像网络结构的小目标检测区域推荐搜索算法。该算法根据区域间多样性颜色直方图相似性构建图像与图的映射关系, 通过图中连通子图的生成获取小目标可能区域。能在生成较少候选区的情况下满足较高的召回率, 进一步优化小目标检测的时间消耗。

关键词 机器视觉; 区域推荐搜索; Girvan-Newman 分裂; 小目标检测

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/AOS201838.0915005

An Algorithm of Small Object Detection Region Proposal Search Based on GN Splitting

Zhao Peiran^{**}, Wu Xinyuan, Tang Xinyu, Shen Xiaohai,

Xu Haiyan, Li Min, Zhang Xuewu^{*}

College of Internet of Things Engineering, Hohai University, Changzhou, Jiangsu 213022, China

Abstract The region proposal search is one of the most active research topics of machine vision. The low efficiency of object detection using the traditional exhaustive search can be improved by optimizing the accuracy of search algorithm. The Girvan-Newman (GN) splitting for community discovery in complex networks is introduced as well as the features of small object regions. A novel method to generate small object regions is proposed by using the network structure of image. The algorithm constructs the relationship between the images and the graphs based on the similarity of color histograms between regions. It can obtain possible regions through the generation of connected subgraphs. This algorithm can meet the higher recall rate in the case of generating fewer candidate regions and further optimize the time consumption of small object detection.

Key words machine vision; region proposal search; Girvan-Newman splitting; small object detection

OCIS codes 150.1135; 100.3008; 100.5010; 110.2960

1 引 言

目标检测^[1]是机器学习中的关键领域。随着深度学习^[2]的不断发展, 有关目标检测的研究日益增多, 在准确率方面也有了极大的提升^[3]。目标检测主要分为候选区域生成^[4]和区域内目标识别两个步骤。前者旨在从图像筛选出目标可能的区域, 后者则负责从区域中判别目标的存在与类别。传统的候选区域生成采用穷举式的滑窗算法, 利用多尺度的窗口遍历整幅图像, 生成大量的候选区域。早期目标检测的研究大多采用这种方法, 实现了诸如行人

检测、人脸检测等系统性的目标检测应用^[5-6], 但是这种方法需要根据待测目标, 人工设计适合目标的多尺度窗口。这种穷举式的算法会产生数量庞大的窗口, 极大地影响检测效率, 难以结合如卷积神经网络(CNN)这种需要大规模运算量的目标识别算法。因此, 研究者们针对穷举式滑窗检测速率较低的问题, 提出了基于区域推荐的搜索方法。该方法旨在尽可能保证召回率的情况下, 利用区域之间的差异性, 生成较少的区域建议窗口, 提高检测速度。Objectness 算法^[7]是最早提出的区域推荐算法之一, 根据颜色、边缘、位置和大小等指标, 提取出图像

收稿日期: 2018-04-09; **修回日期:** 2018-04-25; **录用日期:** 2018-05-02

基金项目: 国家自然科学基金(61671202, 61573128)、国家重点研发计划(2016YFC0401606)

* E-mail: lab_112@126.com; ** E-mail: zhaopeiran1994@126.com

“突出”的区域。基于区域的卷积神经网络(RCNN)算法^[8]采用的是选择性搜索(SS)算法^[9],该方法通过产生预分割区域,利用区域的多样性相似度对预分割区域进行合并,最终生成可能的目标区域;随着对于RCNN的深入研究,SS成为制约其检测速度的主要因素,随后产生了区域推荐网络^[10]。区域推荐网络利用机器学习的方式对目标可能出现的位置和简单特征进行学习,并与后续深度网络共享特征,实现高速的深度检测算法,即基于区域的高速神经网络(Faster RCNN)算法^[10];文献[11]针对车载监控等场景的固有特征,利用在线训练机制,实现了行人检测候选框的快速生成(OL_GMPG)算法;Cheng等^[12]通过训练基于边缘特征的线性分类器,应用滑动窗口的方式快速生成多个候选区域;Zitnick等^[13]利用目标边界的判别,利用非极大抑制等方法动态调整重叠阈值,提高区域生成的召回率。

上述方法设计初衷是为了生成适用于不同尺度的目标候选区窗口,但是在小目标区域提取中存在无用区域过多导致后续检测速率较差的问题。本文提出基于Girvan-Newman(GN)分裂的小目标区域推荐(GN_RP)算法,针对小目标图像的特征,利用图论中图和图像之间良好的映射关系,将复杂网络中用于社区结构的GN分裂算法引入区域推荐搜索中,通过图论中连通子图的概念生成候选区。在VOC2007图像数据库上筛选了部分小目标图像,应用该算法进行处理,取得了84%的召回率,并优化了时间消耗。

2 算法原理

2.1 图像与图的映射

将图像映射为无向加权图的结构再进行处理是一种常见图像处理方法^[14-16]。图中的顶点可以表示为图像中的像素值、像素点坐标或图像中的区域。连接两个顶点的边则表示该顶点对存在某种联系,而边的权重则代表联系的强弱,如两个像素颜色的相似性或两个区域纹理的不相似性。

本文所述的GN_RP算法就是在图的网络结构基础上,利用GN算法生成小目标候选区域。因此需要根据图像与图之间的映射关系,生成无向加权图 $G(V, E)$ 。首先利用分割算法,产生预分割区域集合 $R = \{r_i\}$ 。每一个区域 r_i 代表图 G 的一个顶点 $v_i \in V$ 。若两个区域之间存在联系 $F(r_i, r_j)$,则构成图的一条边 $e_{ij} \in E$,在这里定义,只有在原始图像中存在直接连通的区域对 (r_i, r_j) 存在边 e_{ij} ,

无方向区分。并且区域对 (r_i, r_j) 之间有且仅有一条连接它们的边 e_{ij} 。每条边存在边的权重 $w_{ij} \in W$,权重 w_{ij} 代表了区域对 (r_i, r_j) 之间的相似程度,相似程度越高,连接它们的边的权重则越大。

2.2 GN_RP 算法原理

GN分裂算法^[17]是一种运用于复杂网络中社区发现的经典聚类算法,其基本思想是在构建好的复杂网络中,计算网络中连接相邻顶点的边的边介数^[17],通过不断删除网络中具有最大边介数的边,从整个网络中划分出独立的社团。GN分裂算法为区分一个社区的内部连接和社区与社区之间的连接提供了一种有效度量^[18]。边介数是将两个独立的社团从整体网络中分裂,从而形成两个单独聚类的有效评价指标。只要找到最大边介数边连接的两个顶点,就可以认为这两个顶点分别属于两个不同的类别。一旦将这条边从网络中删除,则两个顶点不再有直接联系。不断重复执行该算法至停止条件,最终所有独立的社团将被明显地划分出来。

若将复杂网络中独立的社团类比为图像中可能的目标区域,利用GN算法的思想,每次从网络中找到权值最小的边,并从图中删除。若一个顶点与周边顶点间的相似度都很小,那么这个顶点与周边顶点的连接都会被删除,从而被分割成一个连通子图。从图像中来看,目标所在区域与其周围背景之间的相似度总是低于目标内部和背景内部的相似度。边的权重越小,代表着这条边连接的区域越有可能属于两个不同的类别(目标与背景)。当图中的边被删除,导致连通子图个数发生变化时,就会提取出一个目标可能的区域。如图1所示,黑色区域与其周边区域的相似度小于黑色区域内部或周边区域内部的相似度。执行GN_RP算法后,黑色区域与周边的区域之间边的联系被删除,黑色区域内部的联系仍被保留,当前图中的连通子图个数为2,黑色区域为所提取出的可能的目标候选区域。

当算法执行至停止条件时,计算获得所有的连通子图集合 G' ,从中删除极大连通子图 g_{\max}' ,即可获得所有的建议区域。极大连通子图 g_{\max}' 在图中表示为 G' 中包含顶点个数最多的子图,在图像中则表示为未被删除连接的初始分割区域集合,一般情况下可以认为该集合代表的区域为图像背景。

3 算法实现

3.1 图像预分割

GN_RP采用快速漂移分割算法(Quick

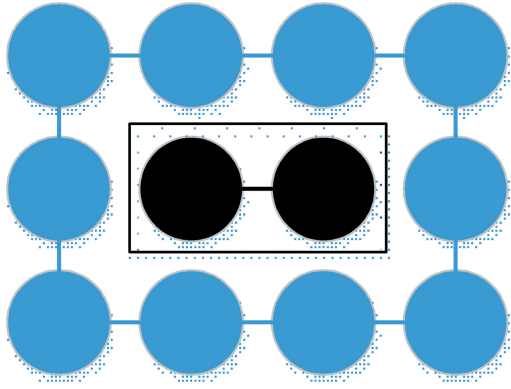


图 1 连通子图示意图

Fig. 1 Schematic of connected graph

Shift)^[19]产生预分割区域。Quick Shift 是一种基于模式搜索聚类的图像分割算法,是在 Mean Shift^[20]和 Medoid Shift^[21]的基础上,通过对判别条件的限制,直接计算而非多次迭代实现的快速分割算法。

设数据点 $x_1, x_2, \dots, x_n \in X, x$ 的 Parzen 密度 $P(x)$ 估算式为

$$P(x) = \frac{1}{n} \sum_{i=1}^N k(x - x_i), \quad (1)$$

式中 $k(x)$ 为任意窗函数。模式搜索聚类算法的核心思想就是将每个点 x_i 划分至不同的模态 $P(x)$, 每一个模态代表不同的类别。Quick Shift 算法通过将点 x_i 移动到最近的具有更高密度的邻域实现模态划分。在算法实现中有关键参数 S_k 和 d_{\max} , 它们分别代表用于平滑样本密度的高斯核大小以及用于数据截止的最大欧氏距离。两者越大,最终被分割得出的区域的个数越少。图 2 为 Quick Shift 分割结果示意图。

3.2 相似度计算

小目标物体在图像中像素占比较小,原始物体中的纹理信息被极大压缩,颜色信息却得到了较大的保留,成为区分不同区域的主要信息。为了充分利用图像信息,本算法利用多样性颜色直方图相似度衡量不同区域之间的相似性^[22]。计算步骤如下:

1) 将单通道灰度划分为 64 个等级,统计生成单通道颜色直方图。

2) 在 Lab、GRB、HSV 和灰度颜色空间,共计 10 个颜色通道上分别构建颜色直方图 $h_i \in H$, 颜色直方图的维度为 $10 \times 64 \times n_k = 640 \times n_k, n_k$ 为单通道中灰度级为 k 的像素个数。

3) 利用巴氏距离计算不同区域,不同颜色空间的直方图的相似性。巴氏距离计算公式为

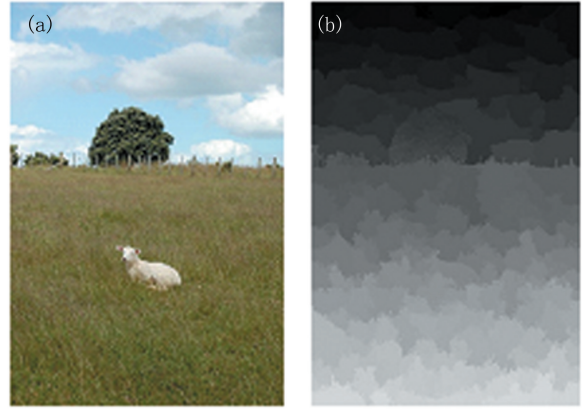


图 2 Quick Shift 算法分割示意图。

(a) 原图; (b) 分割结果图

Fig. 2 Segment by Quick Shift algorithm.

(a) Original image; (b) segmented result

$$\rho(h_i, h_j) = \sum_{i,j=1}^n \sqrt{h(i) \cdot h(j)}. \quad (2)$$

4) 统计不同颜色空间的直方图相似度加权和

$S(i, j) = \sum_m a_m \rho_m$ 作为衡量不同分割区域的相似性指标,即图中边的权重大小。

3.3 算法流程

算法执行流程如表 1 所示。其中对于图的权重进行升序排序是算法在时间上的一个优化。由于原始 GN 分裂算法中边介数在每次删除边之后会发生变化,需重新计算并排序^[18],而 GN_RP 中,一旦区域间的相似性计算完成后不会发生改变。 w_{thr} 为权重阈值只删除所有 $w_k < w_{\text{thr}}$ 的边,一般的可以将 w_{thr} 设置为权重的均值。在计算连通子图时,常用的方法有深度优先搜索或者广度优先搜索。

表 1 算法流程

Table 1 Algorithm flow

Algorithm 1: GN_RP
Input: (color) image
Output: Set of small object location boxes B
Obtain initial regions $R = \{r_i\}$ using Quick Shift
Calculate histograms from different color space in r_i
Foreach Neighbouring region pair (r_i, r_j) do
Calculate similarity $S(i, j)$ as the weight w_{ij}
Generate graph G by neighbor node pair and similarity
Sort w_{ij} in an ascending order
While $w_{ij} < w_{\text{thr}}$ do
Delete e_{ij} from G
Extract set of connected subgraph G' from G_{cut}
Delete the maximum connected subgraph in G'
Extract object location boxes B from G'

4 结果与讨论

4.1 分割算法对区域建议的影响

所有实验均在 Windows 10 系统下利用 anaconda3 进行测试。在 VOC2007 图像数据库上收集部分小目标图像作为测试数据集,待测目标包含车辆、飞机、浮标、船只和各类动物等不同大小和长宽比例的多个类别,每幅图像中可能的目标数未知。所有实验中真阳性 (TP) 需满足重叠度 (IOU) >0.7 。分割算法中关键参数对区域建议窗口影响如表 2 所示。

参数 S_k 和 d_{\max} 实际上控制了图像分割程度,预分割结果对本算法有一定的影响。若分割的太“细碎”,则会导致目标内部被分割成多个部分,目标内

部之间的差异性可能不满足阈值条件,导致目标内部出现多个候选区,IOU 无法满足条件;若分割的太“松散”,则会导致目标与背景在初始分割中被划分为同一区域,无法在后续的分裂中进行区分,候选区域窗口过大。图 3 为部分实验结果。

表 2 分割算法参数对结果的影响

Table 2 Effects of segmentation parameters on the results

Types of different parameters	Recall/%	Number of proposals
$S_k = 5.0, d_{\max} = 20.0$	70.04	12.06
$S_k = 3.5, d_{\max} = 20.0$	71.60	22.17
$S_k = 5.0, d_{\max} = 13.5$	81.26	23.38
$S_k = 3.5, d_{\max} = 13.5$	84.09	41.14
$S_k = 2.5, d_{\max} = 13.5$	78.44	50.49



(a)



(b)



(c)



(d)

图 3 部分实验结果($S_k = 3.5, d_{\max} = 13.5$)。(a)飞机;(b)绵羊;(c)汽车;(d)轮船

Fig. 3 Part of experimental results with $S_k = 3.5, d_{\max} = 13.5$. (a) Airplane; (b) sheep; (c) car; (d) buoy

在区域建议搜索中首先需要关注召回率,若在区域建议中缺失目标,那么在后续检测中,也无法获得目标的准确位置。在设置合理参数情况下,通过预分割算法获得一个既不“松散”也不“细碎”的分割结果,对召回率的提升有一定的帮助。实验中设定的参数对于不同的图像具有一定的普适性,对于简单和复杂背景都有一定的适用性,对于不同类别的

目标无需设置不同参数,能够自适应物体不同的长宽比,并推荐出图像中小目标可能的位置。

图 4 为分割算法产生的区域数量与 GN_RP 算法运行时间的比对。横轴表示 Quick Shift 算法产生的分割区域数量,纵轴表示 GN_RP 运行时间。本文算法在图的生成、排序和连接子图的获取的处理上时间复杂度较低,整体时间消耗与分割区域数

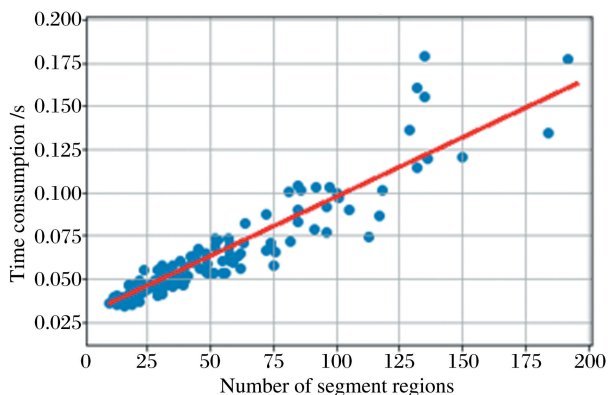


图4 时间消耗与分割区域数量的关系

Fig. 4 Relationship between time consumption and number of segment regions

量呈线性关系。

4.2 算法对比

采用穷举式滑窗搜索和 SS 算法进行对比测试,结果如表 3 所示。时间消耗为生成所有窗口的耗时,统计了所有算法的核心步骤,不考虑预处理时间。

表3 不同算法的结果对比

Table 3 Comparison of results using different algorithms

Types of different search	Recall/%	Time/s	Number of proposals
Exhaustion search (9900)	80.55	0.004	9900
SS($S_{scale}=80, S_{min}=50$)	61.84	0.113	111.86
SS($S_{scale}=30, S_{min}=10$)	82.89	0.565	700.82
GN_RP($S_k=3.5, d_{max}=13.5$)	84.09	0.068	41.14

穷举式滑窗算法根据待测目标预先设置了 3 种不同长宽比和 2 种尺寸大小共计 6 种窗口,每幅图像中共计生成 9900 个候选框。在预先得知目标种类及大致尺寸的情况下,利用人工设计的窗口可以获得较高的召回率,并且生成所有候选区的时间也较少。但是由于穷举式算法生成的候选区数量过多,影响整体检测效率。并且在检测目标更换时,需要根据目标重新设置窗口数据。SS 算法与本文算法类似,需要先利用分割算法生成初始区域,分割的效果也会影响最终召回率。与 SS 算法相比,针对小目标的区域建议搜索这一问题,本文算法能够在生成较少数量的候选区域的同时,提升一定的召回率;生成所有候选框的时间较少,在召回率和时间消耗上具有一定的优势。

实验误差主要来源于阈值的设置。实验中将阈值设置为权重的均值,该阈值不能满足所有的样本,

部分目标与背景间的区分度较低,目标的遗漏导致召回率的下降。此外,可以通过添加如纹理、大小、边缘等多样性相似度或者优化预分割算法的方法进一步优化该算法。

5 结 论

GN_RP 算法根据图像与图的映射关系,利用社交网络中 GN 分裂算法的思想,针对小目标图像特点,设计区域推荐搜索算法,其优点主要包括:1)比穷举式滑窗的召回率高,并且无需根据待测目标的变更设置不同的参数,在提高整体检测结果的召回率同时,优化系统耗时;2)比 SS 算法的召回率高,并且可以通过优化候选区生成时间和生成较少数量的候选区,提高小目标检测效率。

参 考 文 献

- [1] Mo S W, Deng X P, Wang S, *et al.* Moving object detection algorithm based on improved visual background extractor[J]. *Acta Optica Sinica*, 2016, 36(6): 0615001.
莫邵文, 邓新蒲, 王帅, 等. 基于改进视觉背景提取的运动目标检测算法[J]. *光学学报*, 2016, 36(6): 0615001.
- [2] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*, 2015, 521(7553):436-444.
- [3] Xin P, Xu Y L, Tang H, *et al.* Fast airplane detection based on multi-layer feature fusion of fully convolutional networks [J]. *Acta Optica Sinica*, 2018, 38(3): 0315003.
辛鹏, 许悦雷, 唐红, 等. 全卷积网络多层特征融合的飞机快速检测[J]. *光学学报*, 2018, 38(3): 0315003.
- [4] Hosang J, Benenson R, Dollár P, *et al.* What makes for effective detection proposals? [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(4): 814-830.
- [5] Viola P, Jones M J, Snow D. Detecting pedestrians using patterns of motion and appearance[C]// *IEEE International Conference on Computer Vision*, 2003: 734.
- [6] Papageorgiou C P, Oren M, Poggio T. A general framework for object detection [C]// *IEEE International Conference on Computer Vision*, 2002: 555-562.
- [7] Alexe B, Deselaers T, Ferrari V. Measuring the objectness of image windows[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(11): 2189-2202.

- [8] Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation[J]. *Computer Science*, 2013: 580-587.
- [9] Sande K E A V D, Uijlings J R R, Gevers T, *et al.* Segmentation as selective search for object recognition [C] // *IEEE International Conference on Computer Vision*, 2012: 1879-1886.
- [10] Ren S, He K, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks [C] // *International Conference on Neural Information Processing Systems*, 2015: 91-99.
- [11] Qin J, Wang M H. Fast pedestrian proposal generation algorithm using online Gaussian model [J]. *Acta Optica Sinica*, 2016, 36(11): 1115001.
覃剑, 王美华. 采用在线高斯模型的行人检测候选框快速生成方法 [J]. *光学学报*, 2016, 36(11): 1115001.
- [12] Cheng M M, Zhang Z M, Lin W Y, *et al.* BING: binarized normed gradients for objectness estimation at 300 fps [C] // *IEEE Computer Vision and Pattern Recognition*, 2014: 3286-3293.
- [13] Zitnick C L, Dollár P. Edge boxes: locating object proposals from edges [M]. *Chambridge: Springer International Publishing*, 2014, 8693: 391-405.
- [14] Ellis W D. A source book of gestalt psychology [M]. *Michigan: Humanities Press, Psychological Bulletin*, 1939, 36(1): 40-43.
- [15] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation [J]. *International Journal of Computer Vision*, 2004, 59(2): 167-181.
- [16] He Q Y, Li Z L, Wang X Z, *et al.* Automated retinal layer segmentation based on optical coherence tomographic images [J]. *Acta Optica Sinica*, 2016, 36(10): 1011003.
贺琪欲, 李中梁, 王向朝, 等. 基于光学相干层析成像的视网膜图像自动分层方法 [J]. *光学学报*, 2016, 36(10): 1011003.
- [17] Newman M E J, Girvan M. Finding and evaluating community structure in networks [J]. *Physical Review E*, 2004, 69(2): 026113.
- [18] Wang X F. *Complex network theory application* [M]. *Beijing: Tsinghua University Press*, 2006.
汪小帆. *复杂网络理论及其应用* [M]. *北京: 清华大学出版社*, 2006.
- [19] Vedaldi A, Soatto S. Quick shift and kernel methods for mode seeking [C] // *DBLP Computer Vision - ECCV 2008*, 2008: 705-718.
- [20] Cheng Y Z. Mean shift, mode seeking, and clustering [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, 17(8): 790-799.
- [21] Sheikh Y A, Khan E A, Kanade T. Mode-seeking by medoidshifts [C] // *IEEE International Conference on Computer Vision*, 2007: 1-8.
- [22] Wang Q W. *Study on image histogram feature and application* [D]. *Hefei: University of Science and Technology of China*, 2014.
汪启伟. *图像直方图特征及其应用研究* [D]. *合肥: 中国科学技术大学*, 2014.