

基于人类视觉机制的层级偏移式目标检测

秦 升^{1,2}, 张晓林¹, 陈利利¹, 李嘉茂¹

(1. 中国科学院上海微系统与信息技术研究所, 上海 200050; 2. 中国科学院大学, 北京 100049)

摘 要: 当前大多数深度强化学习方法在目标检测方面的召回率较低。为此, 提出一种层级偏移的动态搜索方法。在原有层级搜索的基础上, 采用了锚点的思想, 增加区域偏移, 避免层级产生的区域局限, 使得搜索更加灵活。结合 Double DQN 与 Dueling DQN 的优势, 以 Double Dueling DQN 的网络结构作为智能体深度增强网络的结构。实验结果表明, 与原有层级搜索方式相比, 该方式的目标检测的精确度与召回率较高。

关键词: 人类视觉机制; 深度强化学习; 层级偏移; 目标检测; 马尔科夫决策过程

中文引用格式: 秦 升, 张晓林, 陈利利, 等. 基于人类视觉机制的层级偏移式目标检测[J]. 计算机工程, 2018, 44(6): 253-258.

英文引用格式: QIN Sheng, ZHANG Xiaolin, CHEN Lili, et al. Hierarchical offset object detection based on human visual mechanism[J]. Computer Engineering, 2018, 44(6): 253-258.

Hierarchical Offset Object Detection Based on Human Visual Mechanism

QIN Sheng^{1,2}, ZHANG Xiaolin¹, CHEN Lili¹, LI Jiamao¹

(1. Shanghai Institute of Micro System and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

[Abstract] In order to solve the problem of low recall rate in object detection with the deep reinforcement learning method, on the basis of simulating human visual mechanism, a dynamic searching hierarchical offset method is proposed. It uses the idea of anchors based on the original hierarchical searching method, which adds a region offset. This method avoids the limitations generated by hierarchical searching method, and makes the search more flexible. This paper combines the advantages of Double DQN and Dueling DQN, using Double Dueling DQN network structure as the deep reinforcement learning network of the agent. Experimental results show that the accuracy and recall ratio are higher than the original hierarchical searching method.

[Key words] human visual mechanism; deep reinforcement learning; hierarchical offset; object detection; Markov Decision Making Process (MDP)

DOI: 10.19678/j.issn.1000-3428.0047348

0 概述

从人类视觉注意力机制的角度出发, 人类在定位物体时, 会对图像的整体和部分进行分析, 从图像部件之间的内在联系来进行区域搜索。比如搜索一幅图像上的飞机, 人类一般会先搜索图像的整体, 然后根据大脑中已有的飞机的特征来逐步搜索图像, 对飞机特征最明显的区域搜索, 从而找到飞机。这是一种典型的以目的为导向的目标检测方式。

然而, 现在普遍的目标检测并没有从人类视觉注意力机制出发进行研究。普遍方法是遍历式提取

一系列的独立的感兴趣区域框, 然后再去判断物体。这种方式割裂了图像内部之间的联系, 并没有从人类视觉注意力机制层面对整幅图像和目标进行学习。

因此, 本文提出一种基于人类视觉机制的目标检测方法, 该方法主要基于深度强化学习框架。

1 相关工作

传统的目标检测区域的提取主要有 3 种方式:

1) 区域聚合的方式, 它是通过将相似区域聚合在一起作为物体的感兴趣区域, 例如文献[1-3]; 2) 滑动

基金项目: 上海市科学技术委员会科研计划项目(17YF1427300)。

作者简介: 秦 升(1993—), 女, 硕士研究生, 主研方向为图形图像处理、计算机视觉、模式识别; 张晓林, 研究员; 陈利利、李嘉茂, 助理研究员。

收稿日期: 2017-05-25

修回日期: 2017-06-26

E-mail: qinsheng@shanghaitech.edu.cn

窗口的方式,它是以固定步长,固定大小的感兴趣窗口在整幅图像上进行滑动。例如文献[4];3)基于锚点的方式,它是以不同的点作为中心点,采用不同大小的方框作为基准框,从而提取大量的物体感兴趣区域,例如文献[5-8]。这些区域的提取是相互独立的,分类器给每个区域一个是否包含物体的可信度评分,从而决定这个区域是否是最终的目标区域。然而,这种方式依赖于大量的物体感兴趣区域提取,因此,消耗了大量的资源。

相较于上述的这种目标检测模式,人类视觉机制是一种类似动态决策的过程。它隐含了人类会主动观察显著性目标的方式。这种基于人类视觉注意力机制的相关研究较少,基于传统数学方式而非深度学习方向的视觉注意机制目标检测,例如文献[9-10]速度受限且不具有广泛应用性。目前与本文工作相关的主要有2篇,文献[11]提出主动式物体定位的方法,将物体定位看成一系列的马尔科夫决策过程(Markov Decision Making Process, MDP),使用深度强化学习去训练智能体自主选择感兴趣区域直到智能体找到了最终的物体为止。每一次智能体采取的动作是对当前矩形框进行缩放,改变比例或者是位置。另一个相关的工作是文献[12]提出的层级式目标检测。层级式目标检测定义了一种自顶到下的层级式搜索方式。这种方法的问题在于搜索方式具有局限性,因此,导致提取的很多感兴趣区域与目标物体不能很好贴合,并且单次放缩的失误会带来目标区域的丢失,因此,导致目标检测的召回率比较低。

本文提出了一种新型的层级偏移式搜索方法,在原有层级式的搜索方式基础之上,给出一组偏移搜索组。结合文献[13]提出的 Double DQN 以及文献[14]提出的 Dueling DQN,将2个网络进行融合,使得深度强化学习网络对于动作和当前状态的价值评价更加客观。

2 层级偏移式目标检测模型

本文认为目标检测是一系列的有目标导向的马尔科夫搜索过程。每一步智能体需要进行自主决策关注图像的哪些区域,以最大化搜索过程中的奖励为目标,从而在很少的几步中找到所需的目标物体。

2.1 动作

本文的动作设计分为3组,第1组是层级式的尺度组,每个尺度组有5个动作,每个动作是当前图片子区域的0.75倍。分别对应图片的左上角、右上角、左下角、右下角及图片中间部分。尺度组动作如图1所示。

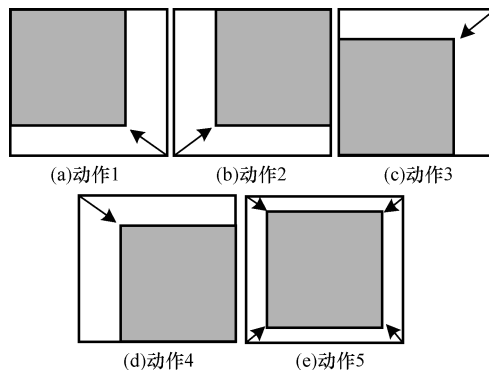


图1 尺度组动作

第2组是偏移组。本文借鉴了锚点的思想,每个偏移组有4个动作,偏移量为图片宽度或高度的0.25倍。分别对应宽度的缩减偏移,扩展偏移以及高度的缩减偏移,扩展偏移。如图2所示,偏移组的目的修正层级组的视野过窄问题,使得视角不会因为单次动作的产生造成之后的视野缺失。

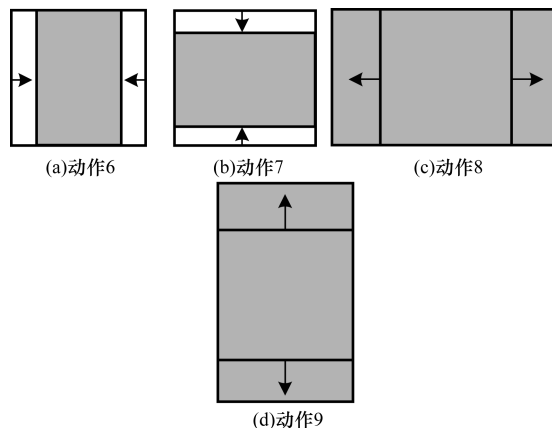


图2 偏移组动作

最后一组为触发组,这个动作代表的是智能体认为当前的感兴趣区域即为物体所在区域,即找到了物体,是整个一系列动作的终结动作。

2.2 状态

对于每一步的动作来说,环境反馈的状态组成包括2个部分:一个是当前步骤产生的图片区域的特征,这个特征是由图片经由文献[15]提出的 VGG-16 的池化层(pool)的第5层 pool5 层输出产生的,图片搜索区域首先变换到统一的 $224 \times 224 \times 3$ 大小,然后把经由 VGG-16 网络输出的卷积特征拉成长度为 25 088 的一维向量作为本文的图片区域特征。经由本文使用的 VGG-16 网络是在 ImageNet 数据集上预训练的模型;另一个是产生的一些列历史动作,它的意义在于使得智能体不局限于当前步骤,使得智能体对于动作的选择具有一个长远的视角。历史动作特征的选择是将一个个动作编码成二进制码,每一个动作被编码成一个10维的向量,只有当智能体选择了当前的动作,当前的动作被编码成1,否则被编码成0。本文采用了4个历史动作作为历史动作序列,因此,

历史动作序列是一个长度为40的一维向量。最后, 本文将这2个特征拼接在一起, 即最终的长度为25 128的一维向量作为本位的状态特征。

2.3 奖励

奖励的设计采用文献[11]中的设计方法。当终止动作没有被触发时, 它的奖励函数公式如式(1)所示。其中, 向量 s 代表采取动作 A 之前的状态, 向量 s' 代表采取动作 A 之后到达的状态。交并比(Intersection-over-Union, IoU)代表搜索区域 r 与目标框 g 的重叠率。具体过程如下, 在 s 状态, 智能体采取一个动作 A , 如果当前采取的动作的所划定的搜索区域 r_t 与目标物体的目标框 g 的重叠率大于之前的采取的动作所划定的搜索区域 r_{t-1} 与目标物体的目标框 g 的重叠率, 那么就给予奖励, 否则给予惩罚, 即 sign 函数给出的+1与-1值。当智能体触发终止动作时, 它的终止动作的奖励函数如式(2)所示, 如果最终触发的动作的区域与目标物体的重叠率大于 τ , 那么认为物体被检测正确, 即给予奖励 η 。在本文的实验中 $\tau=0.5, \eta=3$ 。

$$R_m(s, s') = \text{sign}(\text{IoU}(r_t, g) - \text{IoU}(r_{t-1}, g)) \quad (1)$$

$$R_t(s, s') = \begin{cases} +\eta, & \text{IoU}(r_t, g) \geq \tau \\ -\eta, & \text{otherwise} \end{cases} \quad (2)$$

3 Double Dueling 网络模型

对于整幅图像马尔科夫决策过程的优化本文采用的是深度强化学习网络。具体来说, 文献[16]提出DQN的概念, 进而继续发展为Double DQN, 它的思想是增加一个辅助网络。完全负责产生目标动作的价值, 而主网络负责产生动作。因为被选择的值, 不一定总是最大的值, 这样就避免了被高估的

次优动作总是超过最优的动作, 导致发现不了最好的动作。因此, 学习目标的参数更新可表示为:

$$\theta_t + 1 = \theta_t + \alpha(Y_t^Q - Q(S_t, A_t; \theta_t)) \nabla_{\theta_t} Q(S_t, A_t; \theta_t)$$

$$Y_t^Q = R_t + 1 + \gamma Q_{\text{target}}(S_{t+1}, \arg\max_a Q_{\text{main}}(S_{t+1}, a; \theta_t); \theta_t) \quad (3)$$

其中, 向量 θ_t 代表 t 时刻网络所需要更新的参数, α 代表网络学习速率, 向量 Q 代表辅助网络产生的每一个动作的价值, 它的长度为10, 对应10个不同的动作。向量 Y_t^Q 即 t 时刻对应产生向量 Q 的辅助网络的标签, 由于是非监督学习, 标签的产生由网络自主学习优化产生, 每次的标签由奖励机制模型产生的奖励向量 R_{t+1} 与辅助网络产生的下一个状态的价值组成, 其中, γ 为贴现系数。

进一步, Dueling DQN的思想是让网络分别计算环境本身的价值和选择动作所带来的奖励。在网络的最后部份将环境带来的价值与选择动作带来的奖励进行相加, 这样做的目的是让智能体的学习目标更加明确。如果当前期望价值主要由环境决定, 那么环境的价值就会比较大, 而动作带来的奖励就会比较小; 如果当前的期望价值主要由动作部分决定, 那么环境的价值就会比较小, 动作带来的奖励就会比较大, 分解这2个部分会让学习目标更稳定, 更精确。为了更好地利用这2个网络的特性, 因此, 采用Double Dueling进行决策过程的学习, 从而使得智能体既不会在学习过程中陷入局部最优值, 也会使得智能体学习的目标更加明确。本文的Double Dueling网络结构如图3所示。

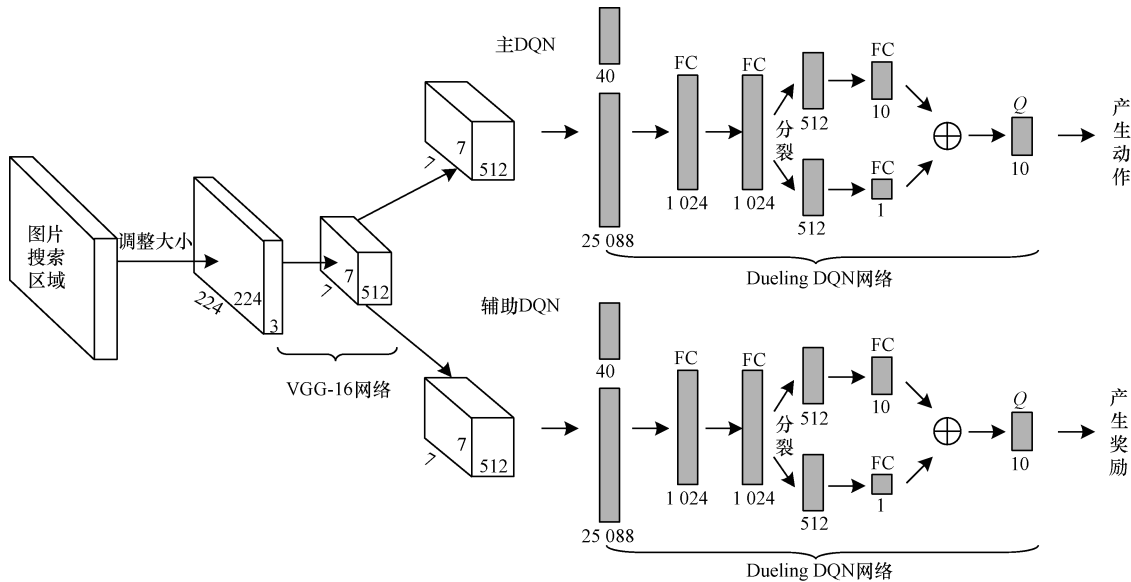


图3 Double Dueling 网络结构

状态特征分别输入2个Dueling网络训练, 在本文实验中, Dueling网络由4个全连接层组成, 共享部分为2个1 024个神经元的全连接层, 然后将得到

的特征分解成两部分分别输入动作主导的价值网络和状态主导的价值网络中, 即10个神经元的全连接层和1个神经元的全连接层。最后把2个价值相加

得到最终价值,主网络用于选择动作,而辅助网络给出奖励用于奖励目标函数的学习。

4 实验结果与分析

4.1 训练方式

这个部分将阐明用于训练 Double Dueling 深度强化学习网络的具体的步骤。

4.1.1 探索与贪婪机制

为了训练 Double Dueling 深度强化学习网络,本文使用一种 ϵ -贪婪策略。它首先以 ϵ 为 1 的概率随机选择动作,然后每个训练阶段衰减 0.1 直到最后随机选择概率衰减为 0.1,之后一直保持 0.1 随机选择的概率。这样做的目的是为了防止智能体在训练的时候陷入局部最优值。同时,为了加速训练,如果当前动作产生的区域与目标物体的重叠率大于 0.5,那么强制使下一次动作为终止动作,即触发态。

4.1.2 智能体学习轨迹

对于智能体的学习,并不能每次给它指定特定的物体让它学习。因此,每一步智能体将会选择与当前搜索区域重叠最大的那个物体进行学习。当然,这种层级偏移式搜索方式中,智能体在学习的过程中可能会改变学习的物体。

4.1.3 经验池

贝尔曼公式如下:

$$Q(s, a) = r + \gamma \max_{a'} Q(s', a') \quad (4)$$

智能体从 (s, a, r, s') 当中逐步学习,而 (s, a, r, s') 可以被叫做以往的经验。因为深度学习需要大量样本,所以传统的 Q-learning 采用的是在线学习的方法,即逐一对新样本学习的方式,但是这对于深度学习并不是适合。为了解决这个问题,采用经验池的学习方式。它的主要思想是,储存一定的经验,并且每次训练的时候随机抽取一部分样本来供给网络学习,这样就能比较稳定地完成学习任务,避免只短视地学习最新接触的样本,而是综合反复地利用过往的大量样本进行学习。本文会创建一个储存经验的综合池,它可以储存一定量的比较新的样本。当容量满后,会用新的样本来代替旧的样本,这可以保证大部分样本有相似的概率被抽取到。在实验中,经验池容量的大小为 1 000,每次强化学习网络选择的样本数为 100。

4.1.4 贴现系数

为了让智能体在一个长期探索的过程中表现良好,不能只看当前动作所带来的价值,而是应该考虑未来所可能采取的动作带来的奖励。因此,使用式(4)的贴现系数 γ ,将 γ 设置为 0.9。

4.2 评价指标与训练参数

4.2.1 评价指标

本文将采用精确度召回率曲线作为评价指标。对于每一次的动作选择都会产生相应的搜索区域,把每一次的搜索区域都看成提取的感兴趣区域。不同于其他的目标检测的置信度评分,本文把每次动作都看作触发动作,即使它不是触发动作,并将强化学习网络产生的向量 Q 值作为某一个区域是否包含物体的置信度。即:

$$Score_A = Q(A_{trigger}) \quad (5)$$

本文采用文献[17]Pascal VOC 挑战中定义的方式,如果当前的搜索框和真实的物体框的重叠率大于 0.5,那么认为是正确的,否则认为是错误的。因此,根据不同的阈值得到不同的精确度与召回率对,从而得到最终的精确度-召回率曲线。

本文平均精确度 (Average Precision, AP) 采用的是 Pascal VOC 2007 的 AP 计算方式,采用 11 点计算法,计算精确度召回率曲线下的面积作为 AP 的结果。

4.2.2 训练参数

本文的学习率设置为 $1e-6$,每个阶段以衰减系数 0.9 进行指数衰减。智能体搜索每个物体的最大步数为 10 步。

4.3 实验结果

4.3.1 精确度召回率曲线对比

图 4 为本文的方法与文献[12]的对比结果。可以明显看出,本文基于 Double Dueling 的层级偏移式搜索方法在精确度和召回率上都有提升。尤其在召回率上,本文提升较大。因此,本文提出的搜索方式更符合人类视觉机制,即从整体到局部,进而从局部进行微调视角避免短视。值得注意的是,精确度在一开始就非常高,这意味着网络对于触发状态的选取更加准确。

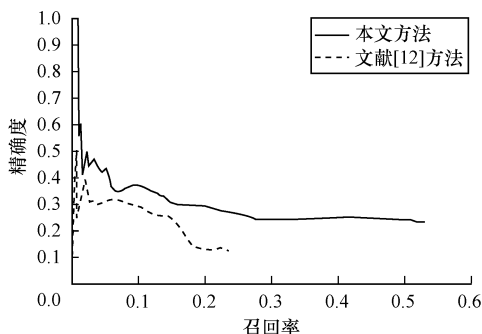


图 4 不同方法的精确度-召回率曲线

4.3.2 AP 对比

本文提出的层级偏移式目标检测的 AP 为 21.9%,而单一层级偏移式目标检测的 AP 为 8.4%。可以明

显看出,本文的方法在AP上有13.5%的提升。同样,由于搜索框比例的限定导致的召回率不高的问题,因此AP的结果相比于传统的目标检测会有一定的差距。

4.3.3 实时性分析

经过时间维度的测试,处理单张图片的平均速度大约在37 ms,即每秒处理大约27张图片,基本满足实时性要求。

4.3.4 不同训练阶段对比

本文同时对不同训练阶段的结果进行了对比。正如所期待的那样,学习的效果是随着训练阶段的增加而变好的,如图5所示。

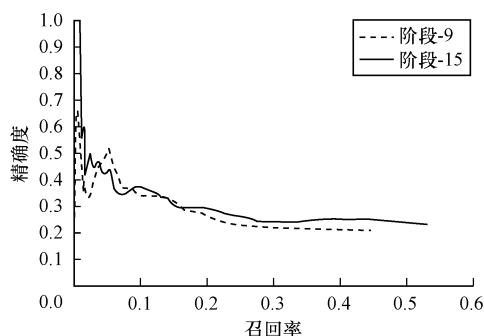


图5 不同阶段的精确度-召回率曲线

4.3.5 定位单个物体所需搜索区域数量分析

图6显示了智能体定位单个物体所需要的区域数量。从直方图可以看出,大部分的物体仅仅只需要搜索一次就可以找到,这意味着对于特征明显或者这个物体占据了图片的大部分区域,智能体能很快定位它。同时可以看出,绝大部分的物体都可以在5步之内被找到。

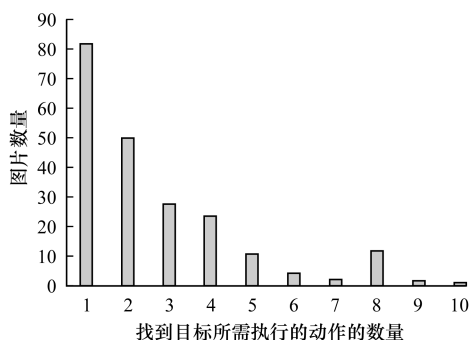


图6 步数直方图

4.3.6 目标检测可视化

图7~图9展示了智能体在图片上进行马尔科夫搜索过程的全部可视化结果。本文设置搜索整幅图像的最大步数为10步,即每张图像产生10个感兴趣搜索框。如图7所示实线框代表搜索过程,即除终止动作之外的动作,虚线框代表终止动作,即智能体认为找到了目标物体,停止搜索。图7展示了

智能体搜索大型目标的过程,可以看出,智能体可以迅速定位目标物体,在1步搜索中确定最终目标物体,十分高效。图8展示了智能体搜索中等目标物体的过程,可以看出,智能体在短短3步之内就确定了最终的目标。图9展示了智能体搜索小型目标的过程,可以看出,智能体主动向特征明显的地方逐步搜索,最终在7步确定了最终的目标物体。

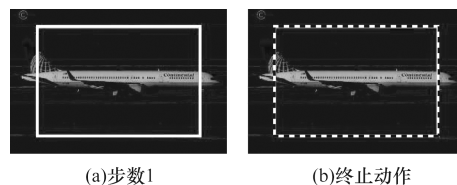


图7 智能体搜索大型目标的可视化过程

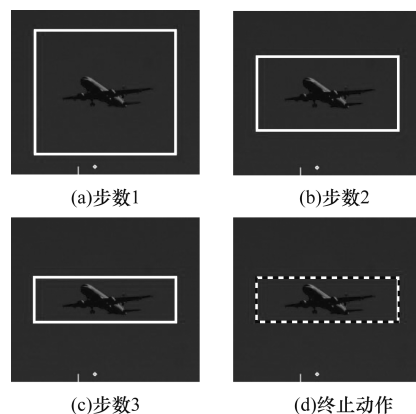


图8 智能体搜索中型目标的可视化过程

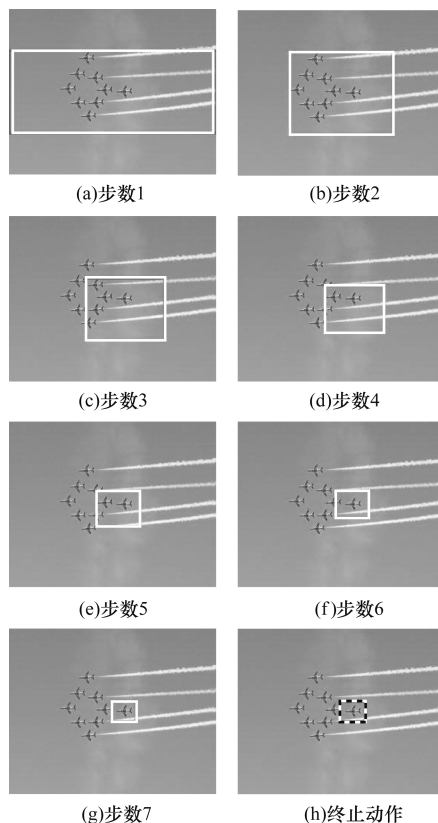


图9 智能体搜索小型目标的可视化过程

5 结束语

本文基于人类视觉机制,提出一种自顶向下带有局部微调的层级偏移式的目标检测方式,并且对深度强化学习网络进行探索。智能体在模拟视觉机制方式的引导下,不断通过自我学习得到图片区域之间的关联性,从而不断优化自我搜索路径。实验结果表明,该方式相较于层级式搜索方式有明显提升。下一步将引入深度信息,探索智能体在三维空间中的搜索路径和方式。

参考文献

- [1] KOEN E A V S, JASPER R R U, THEO G, et al. Segmentation as selective search for object recognition[C]//Proceedings of ICCV'11. Washington D. C., USA: IEEE Press, 2011: 1879-1886.
- [2] CARREIRA J, SMINCHISESCU C. CPMC: automatic object segmentation using constrained parametric min-cuts [J]. Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1312-1328.
- [3] PONT-TUSET J, ARBELAEZ P, BARRON J T, et al. Multiscale combinatorial grouping for image segmentation and object proposal generation [J]. Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(1): 128-140.
- [4] ZITNICK LAWRENCE C, DOLLAR P. Edge boxes: locating object proposals from edges [J]. Journal of European Conference on Computer Vision, 2014, 8693: 391-405.
- [5] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [EB/OL]. [2016-05-09]. https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf.
- [7] ERHAN D, SZEGEDY C, TOSHEV A, et al. Scalable object detection using deep neural networks [C]//Proceedings of CVPR'14. Washington D. C., USA: IEEE Press, 2014: 2147-2154.
- [8] 范琳琳, 冯冬竹, 余航, 等. 自适应轮廓的变分水平集复杂背景多目标检测[J]. 软件学报, 2017, 28(10): 2797-2810.
- [9] 刘芳, 邹琪. 基于视觉注意机制的交通标志检测[J]. 计算机工程, 2013, 39(2): 192-196.
- [10] 朱巨莲, 霍宏, 方涛. 基于自顶向下视觉注意的遥感影像目标检测[J]. 计算机工程, 2011, 37(20): 172-174.
- [11] CAICEDO J C, LAZEBNIK S. Active object localization with deep reinforcement learning [C]//Proceedings of CVPR'15. Washington D. C., USA: IEEE Press, 2015: 2488-2496.
- [12] BELLVER M, NIETO X G, MARQUES F, et al. Hierarchical object detection with deep reinforcement learning [EB/OL]. [2016-11-25]. <https://arxiv.org/pdf/1611.03718.pdf>.
- [13] HASSEKT H V, GUEZ A, SLIVER D. Deep reinforcement learning with double q-learning [EB/OL]. [2016-12-08]. <https://arxiv.org/pdf/1509.06461.pdf>.
- [14] WANG Ziyu, SCHAUL T, HASSELT H V, et al. Dueling network architectures for deep reinforcement learning [EB/OL]. [2016-04-05]. <https://arxiv.org/pdf/1511.06581.pdf>.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2016-09-04]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [16] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing atari with deep reinforcement Learning [EB/OL]. [2016-12-19]. <https://arxiv.org/pdf/1511.06581.pdf>.
- [17] EVERINGHAM M, GOOL L V, WILLIAMS C K, et al. The pascal visual object classes (VOC) Challenge [J]. Journal of Computer Vision, 2010, 88(2): 303-338.

编辑 刘冰

(上接第 252 页)

- [7] 张寒乐. 西林瓶缺陷视觉检测系统研究[D]. 长沙: 湖南大学, 2013.
- [8] KOBAYASHI T. BOF meets HOG: feature extraction based on histograms of oriented p. d. f. gradients for image classification [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE Computer Society, 2013: 747-754.
- [9] NANNI L, MELUCCI M. Combination of projectors, standard texture descriptors and bag of features for classifying images [J]. Neurocomputing, 2016, 173: 1602-1614.
- [10] AL-DMOUR A, ABUHELALAH M. Arabic handwritten word category classification using bag of features [J]. Journal of Theoretical & Applied Information Technology, 2016, 89(2).
- [11] 谢娟英, 王艳娥. 最小方差优化初始聚类中心的 K-means 算法[J]. 计算机工程, 2014, 40(8): 205-211.
- [12] MARIN F B, POTECA U F, MARIN M, et al. Automatic grain size determination in microstructures using computer vision algorithm based on support vector machine (SVM) [J]. Advanced Materials Research, 2017, 1143: 194-199.
- [13] BOUADJENEK N, NEMMOUR H, CHIBANI Y. Writer's gender classification using HOG and LBP features [C]//Proceedings of International Conference on Electrical Engineering & Control Application. Berlin, Germany: Springer, 2017: 317-325.
- [14] 李小林, 周蓬勃, 周明全, 等. 基于可区分二进制局部模式特征的蛾类昆虫识别[J]. 计算机应用与软件, 2016, 33(3): 172-175.
- [15] 徐渊, 许晓亮, 李才年, 等. 结合 SVM 分类器与 HOG 特征提取的行人检测[J]. 计算机工程, 2016, 42(1): 56-60.
- [16] 韩凌波. 一种新的 K-means 最佳聚类数确定方法[J]. 现代计算机, 2013(20): 12-15.
- [17] CORTES C, VAPNIK V. Support vector networks [J]. Machine Learning, 1995, 20(3): 273-297.

编辑 顾逸斐