

一种基于可变形部件模型的快速对象检测算法

李春伟* 于洪涛 李邵梅 卜佑军

(国家数字交换系统工程技术研究中心 郑州 450002)

摘要:为了解决可变形部件模型检测过程中的速度瓶颈问题,该文针对模型的检测流程,提出一种结合快速特征金字塔计算的级联可变形部件模型。由于模型的检测速度主要取决于特征计算以及对象定位这两个过程,提出一种两阶段的加速算法:首先采用尺度上稀疏采样的特征金字塔来近似表示精细采样的多尺度图像特征,以加快特征计算过程;然后在定位过程中结合级联算法,以一个序列模型顺序地评估各个部件,从而快速剪除大部分可能性较小的对象假设,以加快对象定位过程。在 PASCAL VOC 2007 和 INRIA 数据集上的实验结果表明,该算法可以明显加快检测速度,而检测精度仅略有下降。

关键词:快速对象检测;可变形部件模型;特征计算;级联检测

中图分类号: TP391.4

文献标识码: A

文章编号: 1009-5896(2016)11-2864-07

DOI: 10.11999/JEIT160080

Rapid Object Detection Algorithm Based on Deformable Part Models

LI Chunwei YU Hongtao LI Shaomei BU Youjun

(National Digital Switching System Engineering & Technological Research Center, Zhengzhou 450002, China)

Abstract: To solve the speed bottleneck of deformable part models in the detection process, this paper proposes a cascade deformable part model with rapid computation of feature pyramids for the detection process of the model. Because the speed of the detection is mainly determined by the two processes of the feature computation and the object location, a two-stage speedup algorithm is proposed. Firstly, sparsely-sampled feature pyramids on the scale are utilized to approximate finely-sampled multi-scale image features to speed up the process of feature computation. Then combined with the cascade algorithm in the location process, a sequence model is utilized to evaluate individual parts sequentially so as to rapidly prune most object hypotheses of small possibilities in order to speed up the process of object location. The experimental results on PASCAL VOC 2007 dataset and INRIA dataset show that the algorithm in the paper apparently speeds up the speed of detection with minor loss in detection precision.

Key words: Rapid object detection; Deformable part model; Feature computation; Cascade detection

1 引言

对象检测技术是计算机视觉的关键技术和研究热点之一。可变形部件模型^[1](Deformable Part Model, DPM)作为现今最流行的对象检测算法之一,在对象检测领域著名的 PASCAL VOC 竞赛中多次取得优异的成绩,目前已经扩展应用到人体姿态估计^[2],行人检测^[3],人脸检测^[4],以及智能交通^[5]等诸多领域。但是 DPM 模型中涉及的梯度方向直方图^[6](Histogram of Oriented Gradient, HOG)特征金字塔的提取和对象定位两部分的运算量都较大,

导致在单机上处理单张 PASCAL 数据集中的图像时,运行时间在 10 s 左右,无法满足实际应用需求。为此,近年来研究学者针对 DPM 的加速问题开展了系列研究。现有的 DPM 模型加速研究主要集中在以下两个方面:

单类别加速算法:主要关注于单个类别检测中 DPM 模型的加速效果。文献[7,8]通过较早地剪除无望的对象假设来减少计算量,前者根据部件滤波器的得分,使用中间阈值来剪除得分较低的部件假设。但是需要专门的确认数据集来训练中间阈值。后者通过使用多分辨率的层次部件模型,逐次对目标假设进行判断,但是其对于检测精度有一定的影响。文献[9]采用一种类似于文献[7]中的加速策略,通过一种积极的部件选择策略优化了滤波器评估的顺序。文献[10]在所有可能的图像区域上搜索全局最优

收稿日期:2016-01-19; 改回日期:2016-06-08; 网络出版:2016-09-01

*通信作者:李春伟 lichunwei15@126.com

基金项目:国家自然科学基金(61572519, 61521003)

Foundation Items: The National Natural Science Foundation of China (61572519, 61521003)

的分类结果，然而该方法使用最大响应的上确界进行搜索窗口的裁剪操作，其有效性极大地依赖于边界的健壮性。文献[11]采用 GPU 实现并行计算，但是对计算机硬件提出了较高的要求；文献[12]使用快速傅里叶变换来加速检测过程中卷积计算过程。文献[13]通过多种加速方法间的组合来构建完整的加速策略。

多类别加速算法：文献[14]使用大量部件集中的稀疏线性组合来对可变形部件进行编码，以此来压缩模型中所使用有效部件数；文献[15]使用少量基本的部件集合的线性组合来近似表示模板中视角、形变、尺度等方面的变化。类似的方法在文献[16]中也进行了相关研究。这两种方法都是通过学习得到一组共享的基元部件，来获取类别层次的检测效果。文献[17]提出了基于局部敏感的哈希算法来近似 DPM 中的卷积计算，处理的对象类别达到了 100000 种，但是对于检测精度的影响较大。

综上所述，可以看出已有研究中主要关注对象定位加速，并且取得了一些成果，但是对于特征计算加速关注较少。所以想要进一步提高 DPM 模型的处理速度，就不得不考虑特征计算部分的加速问题。由此，本文提出了一种针对可变形部件模型的两阶段加速算法：首先，采用尺度上稀疏采样的特征金字塔近似表示精细采样的多尺度特征金字塔，以此来加快特征计算过程；然后，采用级联算法来加速对象定位过程。为了便于叙述，后续将本文提出的方法称之为结合快速特征金字塔的级联可变形部件模型(Cascade Deformable Part Model with Fast feature pyramid, FCDPM)。

2 基于幂指定律的快速特征金字塔计算

在基于多尺度特征金字塔的滑动窗口模型中为了解决模板尺寸与对象尺寸不一致的问题，需要对目标图像多次进行下采样得到不同尺度下的图像，然后再在逐个尺度的图像上逐个位置上计算图像特征，从而生成精细的特征金字塔，该方法得到的特征精度高，但是相应的时间开销也较大。已有研究人员对特征金字塔进行研究，发现多尺度金字塔尺度间的特征存在一定的关系，且满足幂指定律^[18,19]。由此，本文提出在特征金字塔的计算过程中，利用多尺度特征中的幂指定律有选择地计算部分尺度图像的特征，然后基于这些尺度图像的特征对临近尺度图像的特征进行近似计算，从而减少特征计算的运算量。

2.1 多尺度特征间的幂指定律

为了便于描述，本节先简要介绍一下多尺度图

像间的幂指定律。对于给定的输入图像 I ，定义其对应的通道特征为图像像素的某种线性或是非线性变换的输出响应： $C = \Omega(I)$ 。其中 Ω 表示某种计算函数， C 表示输出响应。定义 f_Ω 为 C 的权值加和： $f_\Omega(I) = \sum_{i,j,s} w_{i,j,s} C(i,j,s)$ ， I_s 为图像 I 在尺度 s 上的重采样图像，其中 I_s 的维数 $h_s \times g_s$ ($h_s = \sqrt{s}h$, $g_s = \sqrt{s}g$) 是图像 I 维数 $h \times g$ 的 s 倍。那么 $f_\Omega(I_s)$ 就被定义为

$$f_\Omega(I_s) \equiv \frac{1}{h_s g_s s} \sum_{ijs} C_s(i,j,s) \quad (1)$$

其中， $C_s = \Omega(I_s)$ 。由此， $f_\Omega(I_s)$ 说明了图像在位置 (i,j) 以及尺度 s 上计算得到的 C_s 的全局平均值。在 DPM 模型中， Ω 表示 HOG 特征映射函数。最终目的是为了得出 $f_\Omega(I)$ 与尺度 s 之间的函数关系。由此设 $\Phi(I)$ 定义了图像 I 的统计信息， $E[\cdot]$ 定义了其在图片集中的数学期望。由文献[19]可知， $E[\Phi(I_{s1})]/E[\Phi(I_{s2})]$ 之值仅与尺度比 $s1/s2$ 相关，独立于尺度值 $s1$ 以及 $s2$ 。由此得到

$$E[\Phi(I_{s1})]/E[\Phi(I_{s2})] = (s1/s2)^{-\lambda_\Phi} \quad (2)$$

其中， λ_Φ 表示对应的拟合参数。针对 HOG 特征计算，使用 $\Phi(I_s) = f_\Omega(I_s)$, $\lambda_\Phi = \lambda_\Omega$ ，由此期望得出图像特征期望与图像尺度之间的关系。

将一幅图像 I 视为 K 幅较小的图像 I^1, I^2, \dots, I^K 的并集，可以得到 $I = [I^1, I^2, \dots, I^K]$ 。由于 HOG 特征对于旋转以及平移并不敏感，因此 $\Omega(I) = \Omega([I^1, I^2, \dots, I^K]) \approx [\Omega(I^1), \Omega(I^2), \dots, \Omega(I^K)]$ ，将结果代入式(1)中得到

$$f_\Omega(I) \approx \sum f_\Omega(I^k)/K \quad (3)$$

此时，如果将 I^1, I^2, \dots, I^K 视为一个较小的图像集合，那么 $f_\Omega(I) \approx E[f_\Omega(I^k)]$ 就表示在这个集合中的数学期望，即图像的全局特征均值 $f_\Omega(I)$ 是子图像集样本特征均值 $f_\Omega(I^k)$ 的数学期望。由于 $f_\Omega(I^k)$ 可以视为较小图像的统计信息并且 $E[f_\Omega(I^k)]$ 表示了在这一较小的图像集合中的期望，那么根据式(2)可以得出， $E[f_\Omega(I_{s1}^k)]/E[f_\Omega(I_{s2}^k)] = (s1/s2)^{-\lambda_\Omega}$ 。由于 $f_\Omega(I_{s1}) \approx E[f_\Omega(I_{s1}^k)]$ 并且 $f_\Omega(I_{s2}) \approx E[f_\Omega(I_{s2}^k)]$ ，进而推断出：

$$f_\Omega(I_{s1})/f_\Omega(I_{s2}) \approx (s1/s2)^{-\lambda_\Omega} \quad (4)$$

2.2 基于幂指定律的尺度间 HOG 特征近似计算

本节将上述定律应用到 DPM 模型中的 HOG 特征计算中。在传统 DPM 模型中，计算尺度图像 I_s 的特征 $C_s = \Omega(I_s)$ 的方法为 $C_s = \Omega(R(I, s))$ ，其中 R 表示采样函数，即需要首先将图像 I 缩放至尺度 s ，然后再计算图像 I_s 的 HOG 特征。根据上述定律， $f_\Omega(I)$ 和 $f_\Omega(I_s)$ 之间存在对应关系，为此在已知原始

尺寸图像 I 的特征 $C = \Omega(I)$ 的情况下, 可以根据其对 C_s 进行近似计算, 从而提高特征计算的速度。

具体过程为, 对于式(4)中设置 $s_1 = s, s_2 = 1$, 由此得到 $f_\Omega(I_s) \approx f_\Omega(I) s^{-\lambda_\Omega}$ 。这一结论不仅适用于原始图像, 对于图像 I 中的采样窗口 w 以及重采样图像 I_s 中的 w_s 都是适用的。由此得到: $f_\Omega(I_s^{w_s}) \approx f_\Omega(I^w) s^{-\lambda_\Omega}$ 。将这一结果代入式(1)中, 得到

$$\frac{1}{|w_s|} \sum_{i,j \in w_s} C_s(i,j) \approx \frac{1}{|w|} \sum_{i,j \in w} C(i,j) s^{-\lambda_\Omega} \quad (5)$$

因为 $\sum_{w_s} C'/|w_s| \approx \sum_w C'/|w|$ 得到 $C' \approx R(C,s)$, 继而得出:

$$C_s \approx R(C,s) \cdot s^{-\lambda_\Omega} \quad (6)$$

2.3 基于幂指定律的 HOG 特征金字塔计算

在 DPM 模型中, 常规的特征金字塔计算方法是在每一个尺度 s 上计算 $C_s = \Omega(R(I,s))$, 从原始图像开始, 在每一个倍频程中相对于上一步以 $2^{1/t}$ 的倍数进行下采样, 直至经历 t 次迭代操作达到原图像的 $1/2$ 为止, 不断重复上述步骤直至检测图像缩小至 5×5 个 HOG 特征块为止。模型中每一个倍频程有 10(即 $t = 10$) 个尺度, 那么在一个倍频程中为了计算多尺度特征, 尺度特征需要重复计算 10 次之多。将上一小节提出的尺度间 HOG 特征近似计算推广到 HOG 特征金字塔的计算之中, 首先计算尺度上稀疏采样的特征金字塔, 依据平均间隔选择若干个基准尺度(例如在每一个倍频程, 即在每 10 个尺度中选择一个基准尺度 $s' \in \{1, 1/2, 1/4, \dots\}$) 计算 $C_{s'} = \Omega(R(I,s'))$, 余下的中间尺度则使用 $C_s \approx R(C_{s'}, s/s')(s/s')^{-\lambda_\Omega}$ 近似计算得到。

3 结合级联算法的可变形部件模型对象定位

3.1 可变形部件模型中对象定位流程

本节首先简要介绍可变形部件模型中对象定位的处理流程。DPM 模型为一个两层模型, 其中根滤波器 w_0 用于检测对象的整体轮廓, n 个部件相对于根滤波器允许有一定的位移, 用于描述目标对象的细节特征及其位置偏移导致的形变特征, 从而模型在处理大尺度外观变化时有优势。第 i 个部件由滤波器 w_i ($i = 1, 2, \dots, n$) 以及形变模型 d_i ($i = 1, 2, \dots, n$) 进行参数化表示。一个对象假设由 $\{p_0, p_1, \dots, p_n\}$ 进行定义。其中 p_0 表示根滤波器的位置, p_i 表示第 i 个部件的位置。根滤波器以及部件滤波器之间通过图结构模型进行连接, 形变模型用于描述两者之间的形变特征。为了检测出图像中目标对象的位置, 在每一个根位置上根据最有可能的部件配置来计算对象假设的整体得分 $\text{score}(p_0)$:

$$\text{score}(p_0) = \max_{p_1, \dots, p_n} \text{score}(p_0, p_1, \dots, p_n) \quad (7)$$

$$\begin{aligned} \text{score}(p_0, p_1, \dots, p_n) \\ = w_0^T \phi_\alpha(p_0, \mathbf{H}) + \sum_{i=1}^n w_i^T \phi_\alpha(p_i, \mathbf{H}) \\ - d_i^T \phi_d(p_i, p_0) + b \end{aligned} \quad (8)$$

其中, \mathbf{H} 表示计算得到的目标图像的 HOG 特征金字塔, ϕ_α 表示相应位置的 HOG 特征, ϕ_d 表示部件与根滤波器之间的形变特征。 b 是为了区分不同的组件, 混合组件的引入是用于表示对象不同的姿态或是形变情况。

对于检测过程中的一个对象假设 $\{p_0, p_1, \dots, p_n\}$, 由于其中仅根位置 p_0 是已知的, 部件位置 p_i 是由最大化部件表现得分减去其相对于根位置发生位移导致的形变开销得到。

$$p_i = \arg \max_p w_i^T \phi_\alpha(p, \mathbf{H}) - d_i^T \phi_d(p, p_0) \quad (9)$$

其中, p 表示部件可能存在的位置。在可能的根位置中根据部件相对于根位置的最佳位置找出总体得分最高的结果, 从而得到一个完整的对象检测结果。

3.2 结合级联的快速对象定位

从上述描述中可以看出, 基于计算出的多尺度 HOG 特征, 在 DPM 模型的对象定位过程中, 需要逐尺度层采用滑窗搜索进行匹配检测。但是, 事实上对象通常仅出现在图像场景中的有限位置, 图像尺度空间上的绝大部分窗口对于视觉对象检测而言是无效的, 如果在每个尺度、每个位置都进行完整的匹配操作, 这就导致了计算复杂度的增加。由此提出在对象定位过程中结合级联检测, 迅速减除大部分可能性较小的对象假设, 以此来提高定位过程的速度。

首先对原模型中部件进行排序来构建层次结构的新模型。在拥有 $n+1$ 个部件(其中包括根滤波器)的 DPM 模型中, 采用一个全局阈值 T 和 $2n(t_1, t_1', t_2, t_2', \dots, t_n, t_n')$ 个中间阈值的序列, 从而得到一个长度为 $n+1$ 的序列模型。在计算匹配得分时, 部件得分是按照一定的次序逐个匹配, 在计算每一个部件表现得分以及相应的变形开销时都要与中间阈值进行判别, 如果小于预设的中间阈值, 相应的目标假设就被剪除。

在每一个根位置 $p_0 \in \mathbf{H}$, 其中进行 n 个阶段(即需要对 n 个部件进行评估)评估得分 $\text{score}(p_0)$, 其中我们所感兴趣的仅是 $\text{score}(p_0) \geq T$ 的根位置。在每个阶段评估 $\text{score}(p_0)$ 时会有两次剪除机会:

假设剪除: 如果在位置 p_0 处前 l ($l = 1, 2, \dots, n$) 个部件的得分小于 t_l 时, 即

$$\mathbf{w}_0^T \phi_\alpha(p_0, \mathbf{H}) + \sum_{i=1}^{l-1} \mathbf{w}_i^T \phi_\alpha(p_i, \mathbf{H}) - \mathbf{d}_i^T \phi_d(p_i, p_0) < t_i \quad (10)$$

算法将会直接剪除在 p_0 处的对象假设，而不再继续评估 $\mathbf{w}_l \sim \mathbf{w}_n$ 部件的得分响应。

形变剪除：在计算部件 \mathbf{w}_l 的得分时，即确定部件 \mathbf{w}_l 的位置 p_l 时，需要在其相对于根位置 p_0 的可能位置 $\delta \in p$ 进行搜索。如果前 l 个部件的得分减去部件 \mathbf{w}_l 相对应的形变开销之后小于中间阈值 t'_l ，即

$$\mathbf{w}_0^T \phi_\alpha(p_0, \mathbf{H}) + \left(\sum_{i=1}^{l-1} \mathbf{w}_i^T \phi_\alpha(p_i, \mathbf{H}) - \mathbf{d}_i^T \phi_d(p_i, p_0) \right) - \mathbf{d}_l^T \phi_d(\delta, p_0) < t'_l \quad (11)$$

算法就会跳过位置 δ 。对于给定的全局阈值 T ，当最终得到的 $\text{score}(p_0)$ 大于 T 时，即得到一次成功的对象检测结果，相应的部件配置结果即为部件位置。

4 结合快速特征金字塔的级联可变形部件模型

本文提出的 FCDPM 算法流程如表 1 所示，其中变量 g 表示每个阶段处理过程中的累计得分， q 表示第 w_l 个部件的得分， m 表示基准尺度的数目， r 表示由基准尺度近似计算的尺度数，检测示意图如图 1 所示。

表 1 FCDPM 算法

输入：检测图像 I ；滤波器 $\{\mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_n\}$ ； 形变模型 $\{\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_n\}$ ；全局阈值 T ； 中间阈值 $(t_1, t'_1, t_2, t'_2, \dots, t_n, t'_n)$ 。
输出：目标对象的位置集合 D 。
(1) FOR $s = s_i (i = 1, \dots, m)$
(2) $C_{s_i} = \Omega(R(I, s_i))$
(3) FOR $s = s_{ik} (k = 1, \dots, r)$
(4) $C_{s_{ik}} \approx R(C_{s_i}, s_{ik}) \cdot s_{ik}^{-\lambda_\Omega}$
(5) END FOR
(6) END FOR
(7) $\mathbf{H} \leftarrow \{(C_{s_1}, C_{s_{11}}, \dots, C_{s_{1r}}), \dots, (C_{s_m}, C_{s_{m1}}, \dots, C_{s_{mr}})\}$
(8) $D \leftarrow \emptyset$
(9) FOR $p_0 \in \mathbf{H}$
(10) $g = \mathbf{w}_0^T \phi_\alpha(p_0, \mathbf{H})$
(11) FOR $i = \{1, \dots, n\}$
(12) IF $g < t_i$ THEN PRUNE p_0
(13) FOR $\delta \in p$
(14) IF $g - \mathbf{d}_i^T \phi_d(\delta, p_0) < t'_i$ THEN PRUNE δ
(15) $q = \max_p \mathbf{w}_i^T \phi_\alpha(p, \mathbf{H}) - \mathbf{d}_i^T \phi_d(p, p_0)$
(16) END FOR
(17) $g = g + q$
(18) END FOR
(19) IF $g \geq T$ THEN $D \leftarrow D \cup \{p_0\}$
(20) END FOR
(21) RETURN D

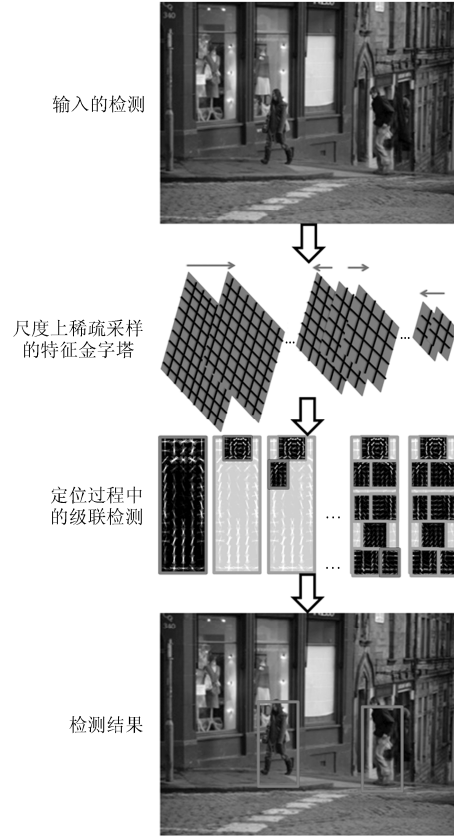


图 1 FCDPM 算法的示例

5 实验结果

5.1 实验准备

为了评估所提算法的性能，本文采用平均精度 (Average Precision, AP) 值作为评价指标，分别在 PASCAL VOC 2007 数据集以及 INRIA 数据集上进行了测试，并且与初始的 DPM 模型以及相关算法进行了比较，AP 值定义如下：

$$\text{AP} = \int \text{PRE}(\text{RE}) d\text{RE} \quad (12)$$

其中，RE 表示召回率，说明了算法的查全性能；PRE 表示精度，说明了算法的检测准确性能。算法的 AP 值通过以下步骤计算得到：(1) 设定不同的检测得分阈值，绘出随着精度递减的召回率/精度曲线；(2) 对曲线下方的区域进行数值积分，得到相应的 AP 值。AP $\in [0, 1]$ 的数值越大，表示算法的性能越好。

5.2 在 PASCAL VOC 2007 数据集上的实验结果

PASCAL VOC 2007 是一个通用的对象检测数据集，其测试集中包含 20 种类别的图像，共有 4952 张测试图片以及 14976 个人工标注的目标对象。通过上面的叙述，设 nApp 为在特征金字塔计算中，由每一个基准尺度所外推计算的尺度数，当 nApp = 0 时，即表示精细采样的特征金子塔。nApp 的取值会影响特征金字塔的计算速度，取值越大，计算速度

越快,但是这也会在一定程度上影响特征金字塔的计算精度,首先讨论了不同的 nApp 取值对于 AP 值以及相对于 DPM 模型检测速度的影响,相应的结果如图 2 所示。

从图 2 可以发现,随着 nApp 值的加大,AP 值总体趋势是不断下降,而相对于 DPM 模型的加速倍数的总体趋势是不断上升。这说明随着 nApp 取值越大,速度越快,但是相应特征金字塔的精度也在下降,从而导致检测精度的下降,这就可以使算法在检测精度和速度之间取得一定的平衡。

表 2 比较了 FCDPM 算法,与初始的 DPM 模型以及若干种加速算法,其中包括 CSC 算法^[7],CF 算法^[8],BB 算法^[10],在 PASCAL2007 数据集上的实验结果,其中 FCDPM 旁的数字表示 nApp 的取值。每一种算法在 20 种类别检测过程中的平均特征计算时间、对象定位时间以及完整检测时间如表 3 所示。

从表 3 可以发现,CSC 算法、CF 算法、BB 算法相比于初始的 DPM 模型在 PASCAL2007 数据集上取得了 10 倍左右的加速效果,本文算法与之相比,可以 3 倍快于这些改进算法,但是平均的 AP 值有一定的下降(特别是与 CSC 算法相比较)。即当 nApp = 9 时,可以取得最快的加速效果,相应的精度下降也最多,但是与 CF 算法和 BB 算法相比,综合考虑精度以及速度,具有较大的优势;而且可以通过调整这一参数,从而在检测精度和速度之间取得一定的平衡。当 nApp = 5 时,平均的 AP 值为 31.3%,当 nApp = 1 时,平均的 AP 值上升到 32.4%,此时与 CSC 算法仅相差 0.4%,检测速度为 CSC 算法的 1.6 倍。

5.3 在 INRIA 数据集上的实验结果

INRIA 数据集是一个常用的行人检测数据集,其中包含 614 幅图像,共有 1237 个正样本和 1218 个负样本。表 4 说明了上述算法在该数据集上的实

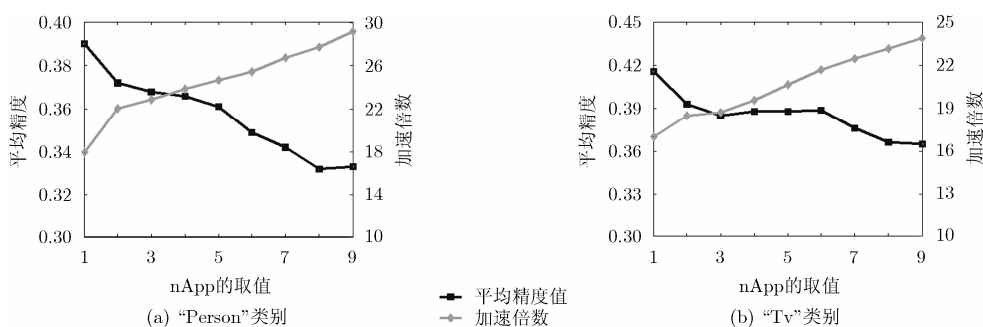


图 2 nApp 的取值对于检测速度以及精度的影响

表 2 不同算法在 PASCAL2007 数据集中 20 个类别的 AP 值结果(%)

算法	Aero	Bicycle	Bird	Boat	Bottle	Bus	Car	Cat	Chair	Cow
DPM	32.1	59.9	10.6	13.5	24.9	47.7	48.9	27.5	17.4	27.0
CSC ^[7]	33.6	58.5	10.2	13.4	25.0	44.7	49.2	28.0	17.1	27.2
CF ^[8]	27.6	58.6	10.5	10.6	16.5	52.7	42.1	27.6	12.9	29.1
BB ^[10]	30.6	53.9	8.7	10.4	22.5	43.2	38.9	17.5	13.5	22.6
FCDPM9	35.9	57.6	8.4	11.8	26.9	44.3	48.1	18.0	14.7	22.8
FCDPM5	35.5	58.6	10.2	11.6	27.4	46.5	50.6	19.4	16.4	23.4
FCDPM1	37.0	59.6	10.3	11.2	27.4	44.3	50.8	20.6	17.9	23.0

Dog	Horse	Mbike	Person	Plant	Sheep	Sofa	Table	Train	Tv	平均
14.4	57.4	50.0	38.4	13.3	21.0	36.8	31.1	47.5	42.3	33.0
13.5	57.7	50.0	38.5	14.1	21.6	36.8	28.6	47.4	42.0	32.8
15.3	55.2	44.6	26.3	12.4	22.7	38.6	32.8	44.8	32.2	30.6
9.0	55.7	47.2	27.4	11.1	17.1	28.3	23.3	44.5	38.4	28.2
8.7	55.3	46.0	33.3	13.3	19.2	29.5	25.9	44.5	36.5	30.0
9.9	53.2	47.0	36.6	15.0	22.8	32.4	26.8	43.1	38.8	31.3
11.6	57.9	50.8	39.0	15.2	23.5	34.4	26.0	47.0	41.6	32.4

验结果。在数据集中的运行时间的比较如表 5 所示。从中可以发现改进算法在行人检测这一领域也取得了在 PASCAL 对象检测数据集中相似的检测结果。FCDPM 算法在该数据集上最快可以取得将近 32 倍的加速效果。当 $n_{App} = 1$ 时，可以取得与 CSC 算法相接近的 AP 值，但是检测速度为 CSC 算法的两倍。从而验证本文算法在行人检测领域也是适用的。

6 结束语

针对可变形部件模型的检测流程，为了加速模

型的检测速度，本文提出了一种两阶段的加速算法，分别采用特征计算中尺度间的幂指定律以及级联算法加速特征计算以及对象定位过程，从而显著地提高模型的检测速度。下一步的工作集中在加速检测过程中的卷积计算，这与之前的工作^[8,11]是互补的，相结合预期可以得到显著的加速效果；还可以与目前流行的对象建议算法^[20]相结合，以进一步提升算法的检测速度。

表 3 在 PASCAL2007 数据集中的平均运行时间(s)

算法	特征计算	对象定位	运行时间之和	加速倍数
DPM	0.376	9.197	9.573	1.00
CSC ^[7]	0.376	0.439	0.815	11.74
CF ^[8]	0.548	0.438	0.986	9.71
BB ^[10]	0.376	1.287	1.663	5.53
FCDPM9	0.090	0.251	0.341	28.10
FCDPM5	0.110	0.275	0.385	24.87
FCDPM1	0.193	0.318	0.511	18.73

表 4 不同算法在 INRIA 数据集上的 AP 值(%)

算法	DPM	CSC	CF	BB	FCDPM9	FCDPM5	FCDPM1
AP 值	88.7	88.6	83.6	85.5	83.9	86.6	87.7

表 5 在 INRIA 数据集中的平均运行时间(s)

算法	特征计算	对象定位	运行时间之和	加速倍数
DPM	0.766	14.425	15.191	1.00
CSC	0.766	0.670	1.441	10.54
CF	1.242	0.445	1.687	9.00
BB	0.766	1.477	2.243	6.77
FCDPM9	0.157	0.323	0.480	31.65
FCDPM5	0.212	0.354	0.566	26.84
FCDPM1	0.326	0.443	0.769	19.75

参考文献

- [1] FELZENSZWALB P, GIRSHICK R, MCALLESTER D, *et al.* Object detection with discriminatively trained part based models[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(9): 1627-1645. doi: 10.1109/TPAMI.2009.167.
- [2] YAO Benjamin, NIE Bruce, LIU Zicheng, *et al.* Animated pose templates for modeling and detecting human actions[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(3): 436-452. doi: 10.1109/TPAMI.2013.144.
- [3] WEN Jia, WANG Xueping, KONG Lingfu, *et al.* Using weighted part model for pedestrian detection in crowded scenes based on image segmentation[J]. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences* 2016, 86(1): 125-136. doi: 10.1007/s40010-015-0231-3.
- [4] OROZCO J, MARTINEZ B, and PANTIC M. Empirical analysis of cascade deformable models for multi-view face detection[J]. *Image and Vision Computing*, 2015, 42(1): 47-61. doi: 10.1016/j.imavis.2015.07.002.

- [5] OHNBAR E and TRIVEDI M M. Learning to detect vehicles by clustering appearance patterns[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(5): 2511–2521. doi: 10.1109/TITS.2015.2409889.
- [6] DALAL N and TRIGGS B. Histograms of oriented gradients for human detection[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005: 886–893. doi: 10.1109/CVPR.2005.177.
- [7] FELZENSZWALB P, GIRSHICK R, and MCALLESTER D. Cascade object detection with deformable part models[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010: 2241–2248. doi: 10.1109/CVPR.2010.5539906.
- [8] PEDERSOLI M, VEDALDI A, GONZALEZ J, *et al.* A coarse-to-fine approach for fast deformable object detection[J]. *Pattern Recognition*, 2015, 48(7): 1844–1853. doi: 10.1016/j.patcog.2014.11.006.
- [9] ZHU Menglong, ATANASOV N, PAPPAS G J, *et al.* Active deformable part models inference[C]. Proceedings of the 13th European Conference on Computer Vision, Zurich, Switzerland, 2014: 281–296. doi: 10.1007/978-3-319-10584-0_19.
- [10] KOKKINOS I. Bounding part scores for rapid detection with deformable part models[C]. Proceedings of the 12th European Conference on Computer Vision, Firenze, Italy, 2012: 41–50. doi: 10.1007/978-3-642-33885-4_5.
- [11] LIU Qi, HUANG Zi, and HU Fuqiao. Accelerating convolution-based detection model on GPU[C]. Proceedings of the IEEE Estimation, Detection and Information Fusion, Harbin, China, 2015: 61–66. doi: 10.1109/ICEDIF.2015.7280163.
- [12] DUBOUT C and FLEURET F. Exact acceleration of linear object detectors[C]. Proceedings of the 12th European Conference on Computer Vision, Firenze, Italy, 2012: 301–311. doi: 10.1007/978-3-642-33712-3_22.
- [13] YAN Junjie, LEI Zhen, WEN Longyin, *et al.* The fastest deformable part model for object detection[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014: 2497–2504. doi: 10.1109/CVPR.2014.320.
- [14] SONG H O, GIRSHICK R, ZICKLER S, *et al.* Generalized sparselet models for real-time multiclass object recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(5): 1001–1012. doi: 10.1109/TPAMI.2014.2353631.
- [15] PIRSIYAVASH H. Steerable part models[C]. Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012: 3226–3233. doi: 10.1109/CVPR.2012.6248058.
- [16] KOKKINOS I. Shufflets: shared mid-level parts for fast object detection[C]. Proceedings of the 14th International Conference on Computer Vision, Sydney, Australia, 2013: 1393–1400. doi: 10.1109/ICCV.2013.176.
- [17] DEAN T, RUZON M, SEGAL M, *et al.* Fast, accurate detection of 100,000 object classes on a single machine[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Portland, USA, 2013: 1814–1821. doi: 10.1109/CVPR.2013.237.
- [18] RUDERM D L. The statistics of natural images[J]. *Network Computation in Neural Systems*, 2009, 5(4): 517–548. doi: 10.1088/0954-898X_5_4_006.
- [19] DOLLAR P, APPEL R, BELONGIE S, *et al.* Fast feature pyramids for object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(8): 1532–1545. doi: 10.1109/TPAMI.2014.2300479.
- [20] HOSANG J, BENENSON R, DOLLAR P, *et al.* What makes for effective detection proposals?[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(4): 814–830. doi: 10.1109/TPAMI.2015.2465908.
- 李春伟: 男, 1991年生, 博士, 研究方向为图像处理、计算机视觉。
- 于洪涛: 男, 1970年生, 研究员, 研究方向为通信、信息系统、社交网络。
- 李邵梅: 女, 1982年生, 讲师, 研究方向为通信、信息系统、图像处理。
- 卜佑军: 男, 1978年生, 讲师, 研究方向为通信、信息系统。