



**本科毕业论文**

|  |
| --- |
| **一种用于通用目标检测的可变形部件模型训练方法** |
|  |

**简舒蕾**

**201230560412**

|  |  |
| --- | --- |
| 指导教师 | **刘昌余 讲师** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 学院名称 |  | **数学与信息学院** | 专业名称 |  | **信息管理与信息系统** |
| 论文提交日期 |  | 2016年4月 19 日 | 论文答辩日期 |  | 2016年5月 5日 |

摘　　　　要

　　目标检测是现今在计算机视觉领域是一个非常活跃的研究方向，在生活中的很多场合中都会用到，比如行人检测、车辆技术、车牌检测等。与目标检测相关的算法有很多，大多是针对特定类别的算法，比如人脸检测。可变形部件模型（Deformable Part Models，DPM ）是目标检测中优秀的算法，自从提出以后，该算法引起了广泛的关注，自提出之后，DPM每年在PASCAL比赛中都取得了很好的成绩。

DPM的目标检测方法是将目标划分为各个部件，各个部件之间通过弹性连接连接方式组合成一个整体，然后检测目标的存在与否。本文在全面细致地研究学习DPM的基础上，训练出基于DPM的通用目标检测模型，同时通过训练出来的模型验证了该算法在各个类别的目标检测效果。在通用目标检测模型训练过程中，通过在ImageNet上下载数据集后，对数据集进行分割制作出正负样本集，根据正负样本训练出的模型对通用目标进行检测。训练出来的模型，大部分模型证实了DPM的有着很好的检测性能。除此之外，由于训练一个模型通常需要花费较长的时间，于是在程序上，提出了使用Parfor进行并行计算代替串行计算，这将大大提高程序的运行速率，节省模型训练时间。

关键字：可变形部件　　通用目标　　数据集分割

**A Deformable Part Models Training Method for Common Target Detection**

Jian Shulei

（College of Information, South China Agricultural University, Guangzhou 510642, China）

**Abstract:** Target detection is a very active research direction in the field of computer vision now. It will be used in many situations in life, such as pedestrian detection, vehicle technology, vehicle license plate detection, etc…There are many algorithms related to the target detection, mostly the algorithms are for the specific categories, such as human face detection. The deformable part models （Deformable Part Models, DPM） is a excellent target detection algorithm, the algorithm caused widespread concern, since it was put forward after, DPM have made good achievements per year in Pascal.

Target Detection based on DPM is divided target into various components. By elastic connections, the connections between the various components combined into a whole, then detecting the target. In the paper, based on the comprehensive and detailed study of learning DPM, the training of the DPM based generic target detection model, through the training of the model to verify the effectiveness of the algorithm in various categories of target detection. In the generic object detection model in the process of training, the ImageNet download data sets, the data set segmentation to produce a set of positive and negative samples, the positive and negative samples are used to train the model according the generic object detection. Trained out of the model, most of the models confirmed that the DPM has a very good detection performance. In addition, because of training a model usually takes a long time, so in the procedure proposed for parallel computation instead of serial computing using Parfor. This will greatly improve program running speed and save the training time of the model.

**Key words:** deformable part common target detection data set partitioning

目 录

[1 绪论 1](#_Toc450089866)

[1.1 研究背景及意义 1](#_Toc450089867)

[1.1.1 VOC挑战赛和DPM 1](#_Toc450089868)

[1.1.2 DPM的特点 1](#_Toc450089869)

[1.2 研究内容及现状 2](#_Toc450089870)

[1.3 论文组织架构 3](#_Toc450089871)

[2 可变形部件模型 4](#_Toc450089872)

[2.1模型定义 4](#_Toc450089873)

[2.2特征提取 5](#_Toc450089874)

[2.2.1像素级别特征提取 6](#_Toc450089875)

[2.2.2像素级别特征的聚集 6](#_Toc450089876)

[2.2.3方向直方图的归一化 7](#_Toc450089877)

[2.3模型训练 7](#_Toc450089878)

[2.3.1参数的学习 8](#_Toc450089879)

[2.3.2模型训练过程 8](#_Toc450089880)

[2.4本章小结 10](#_Toc450089881)

[3 通用目标检测 11](#_Toc450089882)

[3.1通用目标检测训练流程 11](#_Toc450089883)

[3.2基于PARFOR的并行处理 14](#_Toc450089884)

[3.3通用目标检测的特点 16](#_Toc450089885)

[3.4本章小结 16](#_Toc450089886)

[4 实验与结果分析 17](#_Toc450089887)

[4.1实验环境与评价标准 17](#_Toc450089888)

[4.2 基于ImageNet的训练集制作 18](#_Toc450089889)

[4.3实验结果与分析 19](#_Toc450089890)

[4.3.1训练结果分析 19](#_Toc450089891)

[4.3.2训练结果检验 20](#_Toc450089892)

[4.4本章小结 23](#_Toc450089893)

[5 总结与展望 24](#_Toc450089894)

[5.1工作总结 24](#_Toc450089895)

[5.2研究展望 24](#_Toc450089896)

[参考文献 26](#_Toc450089897)

[附录 28](#_Toc450089898)

[致谢 35](#_Toc450089899)

华南农业大学本科生毕业论文成绩评定表

１ 绪论

1.1 研究背景及意义

能否更好更快地从图像（静态或动态）中检测出特定类别的物体是目标检测的关键。比如Google的无人驾驶汽车，其系统要知道准确的道路信息，如果检测到有行人则要迅速做出停车或绕行，这要求系统能快速准确地检测出目标物体。目标检测是计算机视觉的一个基本任务，是应用研究的一个热门领域，近年来各种检测算法不断发展，许多应用也趋于商业化。为了评估和比较各个优秀的目标检测算法的性能，有了VOC挑战赛，在算法的比拼中，目标检测算法性能越来越好，其中一种可变形部件模型检测算法对一般类别的检测能力引人注目，给目标检测算法的发展支出了一个发展方向。

1.1.1 VOC挑战赛和DPM

　　Pascal VOC挑战赛是指Pattern Analysis,Statistical Modeling and Computational Learning for Visual Object Classes，2005年创办至今，VOC挑战赛成为视觉目标类别检测和识别的一个标准。在挑战赛中，它每年为参赛者提供标准化的带有标记信息的图像数据库，并规定了统一的算法性能评估方法（Andrew Zisserman, Christopher K.I.Williams, John Winn et al,2010）。VOC挑战赛大两个主要课题是：一是分类，图像中是否包含一个特定目标类别的任何实例？二是检测，目标实例在图像中什么位置？VOC为参赛者提供了20中的目标类别，包括人、车、狗、沙发等，参赛者可以选择一些目标类别进行处理。在2007年的VOC挑战赛中，芝加哥大学的Felzenszwalb教授及其团队使用了基于可变形部件模型（Deformable Part Models，DPM）的目标检测方法参加了所有目标类别的检测，并在14项类别检测中进入前两名，其中6项获得了第一名，DPM是一个优秀的一般目标类别检测方法。

1.1.2 DPM的特点

第一，DPM使用的特征是方向梯度直方图（Histograms of Oriented Gradients，HOG），该描述子能有效地描述目标的特征。

第二，DPM使用滑动窗口机制，基于滑动窗口的目标检测方法已经成为主流（S Yan, TX Han, X Wang,2009），在行人检测领域，滑动窗口机制的检测器有非常好的性能。

第三，DPM使用隐SVM（Latent SVM）方法训练模型，将目标类别的各个部件看成是隐藏变量，通过挖掘训练数据寻找到最佳的部件位置。

第四，DPM的模型结构是由语法模型（Grammar Models）演化来的混合星形结构语法模型（David McAllester, Pedro F.Felzenszwalb,2011）。该模型将目标实例分解成各个部件，部件之间以弹性连接的方式组合起来成为一个整体。各个部件的检测为整体检测提供了线索，在检测被遮挡的物体时具有明显的优势。

DPM在一般目标检测能力上有着卓越的表现，测试表明DPM在PASCAL（Andrew Zisserman, Christopher K.I.Williams, John Winn et al,2007）、INRIA（N.Dalal）、Caltech（B Schiele,C Wojek,P Dollar et al,2009）等多个数据库上同样具有良好的检测性能。Felzenszwalb团队公开了一个基于DPM的模型训练和目标检测系统，像我们展现了DPM算法实现的各个细节，对于后来者的学习有着重大意义。

本论文将从深入学习DPM出发，以掌握现今主流目标检测技术为目的进行研究学习。通过长时间的学习和应用实践，将基于DPM训练出自己的通用目标检测模型。

1.2 研究内容及现状

通用目标检测就是选取合适的特征来描述目标并利用特征学习适合多类目标的模型，从而检测出场景中的目标和非目标区域。从20世纪70年代开始，许多研究者就在这方面取得一些成果，如光学字符识别（OCR）系统。到了20世纪90年代，学者们把研究重点从简单背景下的目标检测转移到了复杂背景下的目标检测。但当时的算法所建立的模型局限于某一类特定类别的目标物体，如公路路标中使用的圆形特征（ED Micheli, G Piccioli, M Campani,1996），汽车牌检测中使用长宽比特（DW Kim, JK Hang, KK Sang,1996）征等，要用这些算法推广到其它类别的目标检测变得比较困难。于是，到了90年代中后期，学者将注意力转移到非特定类别物体的目标检测算法中。虽然目标检测的相关研究取得一定的进展，但是仍然存在着很大的差距，主要是在通用目标检测中存在着同类别差异、视点变化差异、光照差异、遮挡问题、复杂背景（何祥健等，2006）等挑战。

　　目前，通用目标检测方法主要分为两大类：一类是通过变换或划分得到的特征，另一类是针对目标物体提取出的特征。

第一类是使用较多的类Haar（Haar-like）特征，它是通过变换得到的。在1998年，Papageorgiou等人提出一个目标检测的通用框架（CP Papageorgiou,M Oren,T Poggio,2013），他们从训练集中学习得到Haar小波基特征。由于在算法中用到了上千维的特征向量，算法如果太简单，将会陷入“高维灾难”，于是他们使用了支持向量机（SVM）对上千的特征向量进行降维，得到较少的支持向量。作者通过实验证明了该方法在人脸和行人两个目标类别检测中都能取得不错的效果。

在2001年，Mohan和Papageorgiouz提出将待检测的目标分解成不同的部分（A Mohan, C Papageorgiou, T Poggio,2001）。以行人检测为例，用滑动窗口扫描图像后，分别对头、左上肢、右上肢、腿进行单独训练形成四个SVM分类器。然后将四个值填充到一维向量中，再训练一个总的SVM分类器判断是否对应了一个人。这种基于不同部分的检测方法对变形大的检测目标有较好的检测精度。

第二类是通过目标物体提取出来的特征，如利用纹理或边界特征。Agarwal等人选用Forster算子进行特征的提取（A Awan, R Dan, S Agarwal,2004），然后通过聚类形成不同的特征。Leibe等人使用了Harris算子提取特征（A Leonardis,B Leibe,B Schiele,2004），再经过聚类形成一个集合。相对于Agarwal，Leibe等人不是使用特征之间的简单关系作为检测目标的特性，而是训练一个隐形状模型（implicit shape model）描述目标物体特性。Dalal等人（B Triggs, N Dalal,2005）将HOG和SIFT描述子进行结合，从而得到一种局部区域特征。通过与小波等特征的对比试验证明了HOG 是一组鲁棒性非常好的特征。这种特征很好的描述了检测目标的局部区域特征。Felzenszwalb等人采用了Dalal提出的HOG特征（D Mcallester,D Ramanan,P Felzenszwalb,2008;MA David,PF Felzenszwalb, RB Girshick et al,2010），但不同之处在于该模型是由一个根模板和若干个部件模板组成，即我们所说的可变形部件（DPM）模型。这样既可以表述目标的整体结构，又可以兼顾目标的局部形变。DPM理论获得了巨大成功，不但在之后的 VOC挑战赛中频繁被使用，最新的研究还探索了更多的应用领域。

1.3 论文组织架构

本文共分为五章，每一章节从不同的角度介绍本的相关内容。具体章节安排如下：

第一章绪论，主要介绍研究课题的背景及意义，和现今课题的研究内容和现状。

第二章可变形部件模型，对可变形部件模型从模型的定义到特征的提取，到模型的训练过程深入了解了DPM。

第三章基于DPM的通用目标检测，在第二章的了解之后，第三章基于DPM训练通用目标检测模型，从ImageNet下载数据到分割数据集，到后来的训练和检测，训练出通用目标模型。并提出用并行计算提高程序的运行效率，最后分析通用目标检测特点。

第四章实验与结果分析，介绍实验的一些情况，比如说实验环境、实验对象等；在训练模型前期，展现了样本的具体制作过程；最后对实验结果进行了分析和验证。

第五章总结与展望，对本文进行了总结，并对通用目标检测存在的问题及未来的研究方向提出了一些想法。

2 可变形部件模型

2.1模型定义

可变形部件模型（DPM）是Felzenszwalb等人在2007年提出的一种用于目标检测的可变形部件模型，它是判别训练的、多尺度的。并在2007年的PASCAL VOC物体检测竞赛中获得了冠军，平均精度（Average Precision，AP）比2006年PASCAL人体检测竞赛结果高了两倍，比2007年PASCAL目标检测竞赛中20个类别其中的10个的最优结果还要好。可变形部件模型采用了Dalal和Triggs提出的HOG特征，该模型是由一个根模板和若干个部件模板组成的。这样既保证了目标的整体结构，又能兼顾到部件结构的形变。

DPM模型包括一个根滤波器和若干个部件滤波器。其中根滤波器是粗糙的，而部件滤波器表示了目标的各个部件，是高分辨率的。这里的滤波器是一个权重矩阵，一个*w\*h*大小的滤波器F 是一个含*w\*h*\*9\*4个权重的向量。滤波器的得分或响应是滤波器的权重向量和特征图上相应位置的特征向量的点积。详见图1。

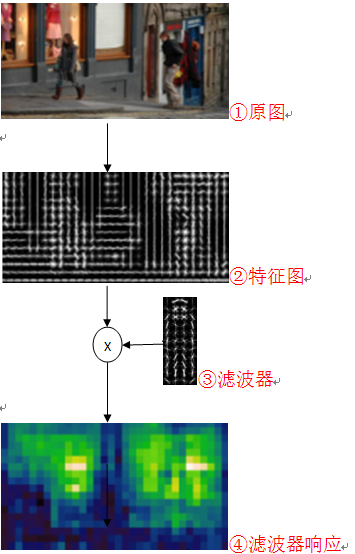


图1 根滤波器处理过程

图1展示的是一个根滤波器的处理过程。图1中将①经过HOG变换后得到②，③是一个大小为*w\*h*的滤波器，②和③的点积得到④，高亮的地方就是目标可能出现的地方。相比于根滤波器的处理，部件滤波器的处理是将特征图放在比根滤波器高两倍的分辨率下。在高分辨率下，可以捕捉到比根滤波器捕捉到更细腻的信息，从而获得目标的部件信息。

DPM模型用（F0，P1,…,Pn,b）来表示，其中F0表示的是根滤波器，而Pi是第i个部件模型，b表示部件滤波器在高分辨率下相对于根滤波器的偏移量。Pi用（Fi,vi,di）进行描述，Fi表示部件滤波器，vi 是未发生形变是部件相对于根部件的标准位置，称为锚点位置。di是检测到的位置相对于锚点位置的偏移。

一个目标的得分等于部件滤波器在各自位置上所得的分数之和减去各个部件的变形花费（即部件位置相当于锚点位置偏移的花费），最后加上偏移量。用z=（P0,…,Pn）表示目标的得分，Pi表示第i个部件模型。用公式表示如公式（2.1）。

（2.1）

其中，

（2.2）

表示第i个部件相对与锚点位置的形变，并且用

（2.3）

来做形变特征。.

一个目标的得分也可以用点积来表示，，其中是模型参数向量，是特征向量，如下公式：

（2.4）

（2.5）

上述几个式子将模型用线性分类器表达出来，在本文将用LSVM对参数进行学习。

2.2特征提取

由于图像目标局部的表象和形状可以用边缘的方向或梯度密度分布来进行描述，所以DPM采用了HOG作为特征描述子。HOG特征提取的具体实现方法首先将图像划分成小的连通区域，称之为cell；然后采集cell中各个像素点的边缘或梯度的方向直方图；最后把直方图组合起来就构成了特征描述器。利用该方法提取边缘和梯度的特征能很好的描述局部形状，由于是局部提取，所以对光学变化有很好的不变性。HOG描述子比现存的特征集有更好的表现。

HOG描述子提取可以分为三个步骤：像素级别特征提取、像素级别特征的聚集、方向直方图的归一化。具体步骤下面逐一介绍。

2.2.1像素级别特征提取

像素级别特征提取为计算每个像素的梯度，包括梯度方向和梯度大小。常用方法是采用如(B Triggs, N Dalal,2005)所述，首先用[-1，0，1]对原始图像做卷积运算，得到x横轴方向的梯度分量；再用[1，0，-1]对原始图像做卷积运算，得到y纵轴方向的梯度分量，然后用公式计算梯度方向和大小。

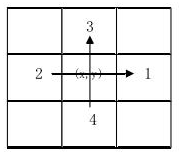


图2 像素梯度的计算

如图2所示，设f为像素点（x,y）处的像素值，梯度大小用r（x,y）表示。

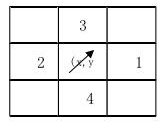


图3 像素梯度特征表示

　　梯度大小的计算是分别计算水平方向和垂直方向的像素差值，dH =f1-f2，dV=f3-f4，如图3所示其梯度方向为θ（x, y），以上两个计算公式分别如（2.6）和（2.7）所示：

（2.6）

（2.7）

2.2.2像素级别特征的聚集

cell是由不重叠的8\*8像素点组成的区域。将cell中梯度方向360度划分为9个方向块，每个像素点梯度方向对应到cell中其中一个方向块。而每个像素点的梯度大小将作为值直接加到该方向上，从而形成方向梯度直方图，如图4所示。

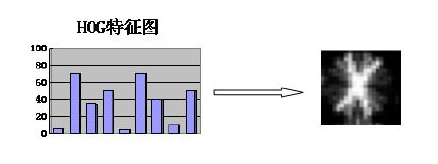


图4 梯度方向直方图

每个cell可以统计出一个一维的梯度方向直方图，这些直方图表示了图像目标的局部特征。

2.2.3方向直方图的归一化

将2\*2个cell组成连通的block，将cell内的特征向量串联起来就组成了block的HOG特征。各个block之间是相互重叠的，所以每个cell内的特征能够多次出现。进行归一化后，由9维变成了36维的HOG描述符。HOG描述符能够进一步对光照、阴影等有很好的不变性。

2.3模型训练

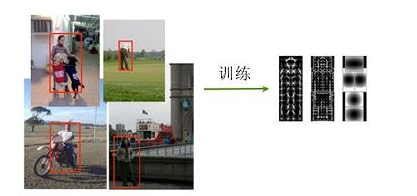


图5 模型训练

　　如图5所示，目标包围盒（Bounding Boxes）标出了图像中目标所在区域，除此之外，图片还标明每个目标的类别（{1，-1}，1表示正样本，-1表示负样本）。通过对这些图片的学习训练，得到包括根模型和部件模型的DPM模型。

2.3.1参数的学习

训练数据集D={（x1,y1）,…,（xm,ym）},yi∈{-1,1}中样本x的分值可用公式（2.8）表示：

（2.8）

　　β是模型参数向量，z是隐藏变量，是使模板与限定框覆盖比率最高的根位置。集合Z（x）定义了样本x所有可能的隐藏变量值。Z确定之后，就是一个标准的SVM训练过程，这一过程由公式（2.9）完成。

（2.9）

　　寻找β使得LD（β）最小，那么就完成模型的训练了。

设C为目标类别，β的学习流程如表1：

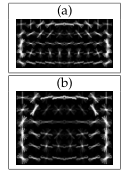
表1 参数学习流程图

|  |  |
| --- | --- |
| **Input:**  正样本P={（I1,B1）,……,（Im,Bm）}，Ii是图片，Bi是标注的包围盒  负样本N={J1,……,Jm}  初始化β,负样本集合Fn = Φ  **Output:**  β | |
| 1: | **for** relabel = 1:num-relabel |
| 2: | 对于正样本，设置正样本集合FP=Φ |
| 3: | **for** i=i：n |
| 4: | 重新标注正样本，将最好的结果加入Fp |
| 5: | **end** |
| 6: | 对于负样本，学习难例 |
| 7: | **for** datamine = 1:num\_datamine |
| 8: | （a）更新负样本集合，加入难例，去掉简单的，如果没有出现新出现负样本或者内存满了则终止。  （b）用新的到的正样本和负样本集合训练β，否则重复（a）。 |
| 9: | **end** |
| 10: | **end** |

2.3.2模型训练过程

（1）初始化根滤波器

下面的训练过程将以汽车为例。如果要训练2个组件（即汽车的正面和侧面）的混合模型，首先将汽车的正样本集P中的包围盒按照长宽比分为P1，P2组，每个组内汽车包围盒具有相似的长宽比。根据2组包围盒每组对应训练出1个汽车根滤波器，共有2个不同的汽车根滤波器F1，F2。用SVM训练Fi。在（I，B）∈Pi中，缩放B中汽车的区域，产生汽车的正样本。从负样本集N中随机选择产生不包含汽车的图像作为负样本。图6（a）中显示的是训练的汽车侧面模型，图6（b）显示的是训练的汽车正面模型。

  
图6 汽车模型的初始根滤波器

（2）合并组件

将2个初始化根滤波器联合形成不含部件的混合汽车模型，在数据集P和N上（不分组不缩放）用2.3.1中训练过程训练联合模型的参数。这个时候组件标识和根位置成为每个样本的隐藏变量了。

（3）初始化部件滤波器

将根滤波器的值插值到两倍空间分辨率，用来初始化部件滤波器。每个汽车组件模型的部件个数固定为6个，放置部件在根滤波器的高能量区域，或者放在根滤波器的中心垂直轴上，或者关于中轴两两对称。一个部件的位置确定后，它所覆盖区域的能量值设为0，然后继续寻找下一个最高能量区域，直到6个部件的位置都确定。

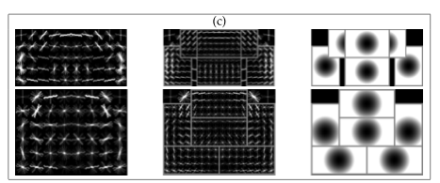


图7 汽车模型的初始部件滤波器

图7（c）显示了初始化部件滤波器的结果。

2.4本章小结

本章介绍了基于DPM模型的物体类检测算法。该算法采用的是HOG特征作为描述符提取底层特征，然后根据模型的定义对模型参数进行学习训练。在对Felzenszwalb提出DPM模型的系统学习后，将在第三章阐述DPM模型在通用目标检测的应用。

3 通用目标检测

论文在学习DPM的基础上，将DPM应用于对通用目标检测上，找到一种用于通用目标检测模型的训练方法。下面将首先介绍通用目标的训练流程。

3.1通用目标检测训练流程

基于DPM，本文将在ImageNet数据集上作为原始数据的来源，训练出通用目标检测模型。图8是通用目标的流程图。本文将以laptop为例进行整个流程的示范。

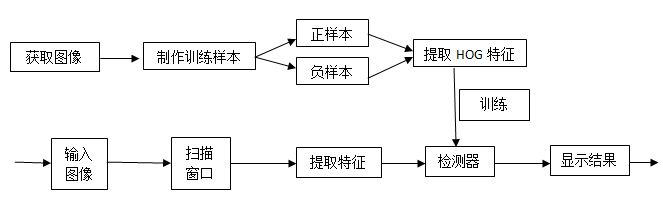


图8 流程图

（1）获取图像

DPM在PASCAL　VOC获得了较好的检测结果，本文将在ImageNet获取所需的图像集。ImageNet是根据模拟人类的识别系统建立的图像识别数据库，目前在世界上是最大的图像识别数据库，能够从图像识别相关物体。本文将在ImageNet上获取10个左右的目标图像，包括牛、蜗牛、飞机、汽车、热气球、蜡烛、小刀、船、手提电脑、足球这十类目标。

图像的具体下载流程如图9所示。

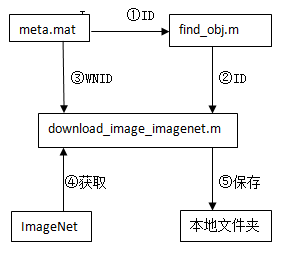


图9 图像下载流程图

在图9可以看到图像下载的流程图，首先了解下meta.mat文件。meta.mat文件为元数据文件，里面定义了图像的一些基本属性，详见图10。

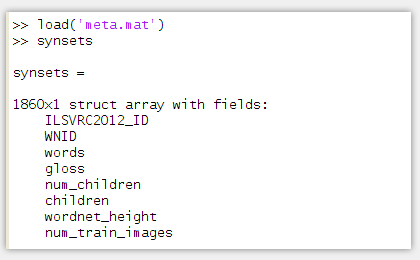


图10 meta.mat

在图10可以看到，meta文件表明了图像的ID、words、孩子等属性。在图9中通过find\_obj.m（见附录A）找到相关目标在meta中的ID值是多少，并返回给find\_obj.m。比如我们需要找出手提电脑的ID号。从图11中可以看出laptop的ID为228。

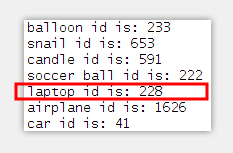


图11 laptop的ID

找到laptop的ID后，在图9的②中，download\_image\_imagenet.m（见附录B）将从find\_obj中获得laptop的ID，根据ID值download\_image\_imagenet.m在③中获得laptop的WNID，如图12所示。

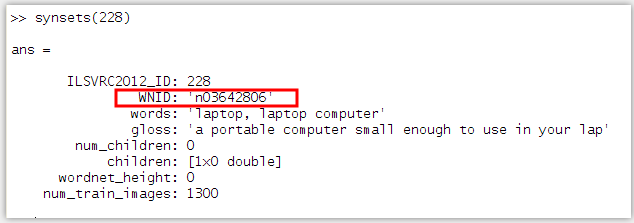


图12 laptop的WNID

获得laptop的WNID后，download\_image\_imagenet.m将根据WNID在ImageNet获取图像相关的信息，将信息进行处理后，将图片保存到本地。如图13是下载完成的laptop图像集。

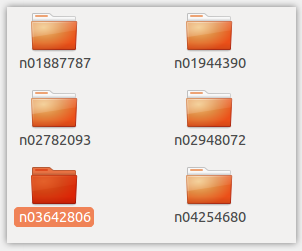


图13 下载完成的laptop图像集

（2）制作训练样本

由于训练的难度和复杂度，每一个类将在下载好的图像中提取出200幅作为该类正样本集。如图14是提取出laptop的200个作为正样本。

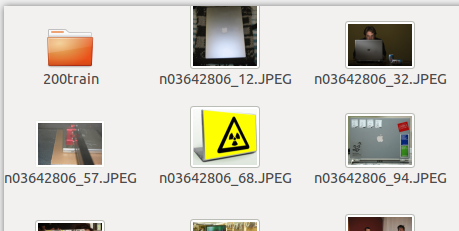


图14 laptop正样本集

（3）提取HOG特征

正负样本制作完成后，将对样本提取HOG特征，具体步骤见2.2小节。

（4）训练

模型的训练过程实际上是一个分类的过程，这里用的是LSVM分类器进行分类。具体见2.3节。

（5）检测匹配过程

模型训练完成后，将对图像进行检测。图15是laptop的训练模型。

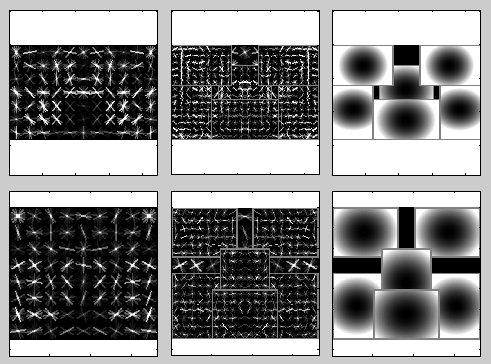


图15 laptop模型

检测过程是首先将原始图像转换成特征图，然后将特征图跟训练好的滤波器模型进行点乘，得到滤波器的得分，高亮的地方就是得分高的地方（详见图1），也是目标最可能存在的地方，从而检测出目标。图16是laptop检测出来的结果。

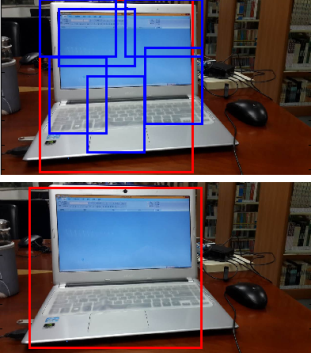


图16 图像检测过程

3.2基于PARFOR的并行处理

在程序运行过程中，消耗计算机资源最多的往往是循环，把循环并行化，可以加快程序的运行速度。并行计算的原理就是将代码分配到多个处理器中进行运算。例如8核的机器，就可以同时调动8个处理器来运算。在Matlab中，提供了Parfor关键字，可以在单台多核机器或集群上实现并行计算，从而加快程序运行的速度。

Matlab的并行计算实质还是主从结构的分布式计算。当你初始化Matlab并行计算环境时，你最初的Matlab进程自动成为主节点，同时初始化多个Matlab计算子节点。Parfor的作用就是让这些子节点同时运行Parfor语句段中的代码。Parfor运行之初，主节点会将Parfor循环程序之外变量传递给计算子节点。子节点运算过程时互不干扰，运算完毕，则应该有相应代码将各子节点得到的结果组合到同一个数组变量中，并返回到Matlab主节点。

根据Parfor写了个小程序测试并行计算与串行计算之间在单核和多核下的运行速度。图17是测试的小程序。

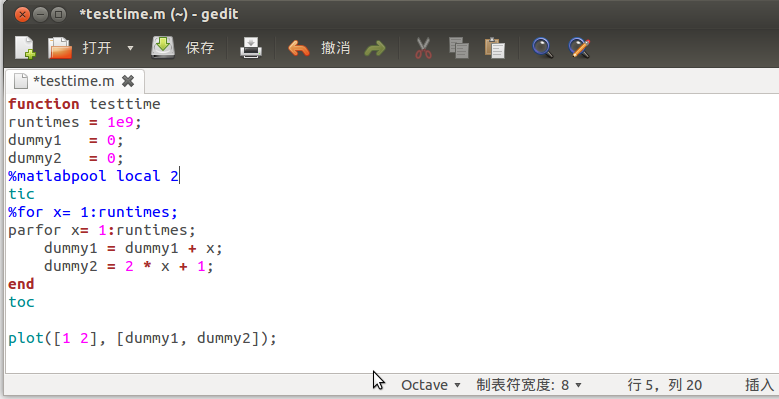


图17 texttime.m

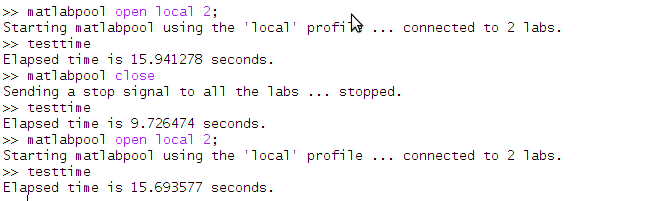


图18 单核下parfor 与for

单核机器下Parfor的开支反而降低了程序的运行速率，无法实现并行计算，那么来看看多核情况下的并行计算的速率是怎样的。详见图19。

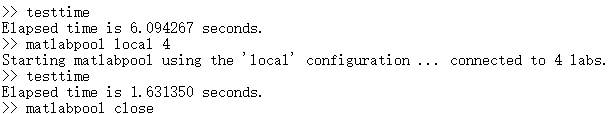


图19 4核下parfor与for

由上面实验可以看出，使用Parfor这个循环进行并行计算速率将是for循环的几倍，如果是4核或者是8核，效果将更好。在通用目标检测模型训练实验中，通常一个模型的训练需要将近10个小时左右，这样对于大规模的训练提出一定的挑战性，提出Parfor这个循环，改进代码，将大大提高代码执行的效率，在现今有大数据的背景下，并行计算将给我们带来更大的效率。

3.3通用目标检测的特点

特定的目标检测算法比如行人检测，人脸识别等都有其针对性的算法，本文侧重于对通用目标检测的检测，相比较于特定目标的检测方法，通用目标检测对算法的要求更高。

首先，模型要足够多以表达丰富的目标类别，要能够在同一个框架下对不同目标对象进行解释。由于时间、资源有限，所以本文选取了10个左右的目标进行模型训练，主要是牛、蜗牛、飞机、汽车、热气球、蜡烛、小刀、船、手提电脑、足球。这些模型中，类别间相似度低，能更好表达目标类别，丰富种类，更好的在同一框架下对不同目标对象进行解释。

其次，模型要选择表示能力强的特征，这样才能描述和检测大部分目标类别。描述图像特征的一般做法是使用所谓的描述子，比如将图像区域中所有的像素加入一维向量进行特征提取后，通过计算卷积相应来根据两个描述子之间的相似性进行判断。这种方法的缺点是描述子向量的维数太高，特征冗余。在之后，研究者分别提出了形状上下文，自旋图、SIFT以及 PCA-SIFT等图像描述子。在2005年，牛津的Mikolajczyk等人比较分析了一些主流的描述子性能，实验发现SIFT描述子在区分行和鲁棒性方面都表现很好。而HOG特征描述子可以视为是SIFT的扩展，性能上比SIFT更优。本文采用的是HOG特征描述子作为特征的表述，在通用目标检测中，能更好的描述大部分的目标类别，更易于通用目标的检测。

最后，模型要包含目标实例的变化，在描述目标实例时有一定的灵活性。

通过对以上特点的分析，相信大家对通用目标检测有更深入的了解了。

3.4本章小结

本章在开始的时候介绍了通用目标检测算法的发展情况之后，根据Felzenszwalb 提出的DPM模型流程上进行一些改进，详细介绍了作者在训练模型过程中做的一些流程，和介绍了通用目标检测的一些特征，让大家对模型的训练过程有了更近一步的了解。

本章基于训练的效率提出了一些改进的方法，比如把串行计算for改进成用Parfor进行并行计算，在程序运行效率上有很大的提高。

4 实验与结果分析

4.1实验环境与评价标准

（1）实验环境

　　算法实现的代码在matlab上编写，在matlab 2012a上运行。

　　测试平台： Ubuntu14.04操作系统，Intel（R） Core（TM） i5-3317U中央处理器，单核，主频1.7GHz，4GB的物理内存。

（2）实验对象

　　实验使用图像来自ImageNet物体检测识别所使用的标准图像库，提取该库训练库有标准图像961幅，验证图库1331幅。

（3）评价标准

在VOC挑战赛中经常用到的一种评估手段是利用平均精度（Average Precision，AP），本文采用AP数值大小来作为评价目标检测准确率的高低。为了了解AP的计算方法，首先定义什么是正确检测：

正确检测值检测结果预测的目标包围盒Bp与实际包围盒Bt的重叠率a超过了50%。

当系统预测结果有多个目标包围盒与实际包围盒重叠的时候，只取其中一个作为正确检测。其中a定义为公式（4.1）

（4.1）

AP由精度/召回率（precision/recall）计算得到，精度表示正确检测到的目标数占总共检测到的目标数；召回率表示正确检测到的目标数占图库中的总目标数。Precision的定义如公式（4.2），recall的定义如公司（4.3）所示。

（4.2）

（4.3）

　　精度和召回率是互相影响的，理想情况下能做到两者都高，但是一般情况下精度高、召回率就低；召回率高、精度低。用不同的召回率可以得到不同的精度。AP 是对不同召回率点上的精度进行平均。对一个有序的列表，计算 AP 的时候要先求出每个召回率位置上的精度，然后对所有召回率位置上的精确度做平均。通过系统在测试集上的精度-召回率曲线的AP对系统进行打分。AP曲线图如图20所示。

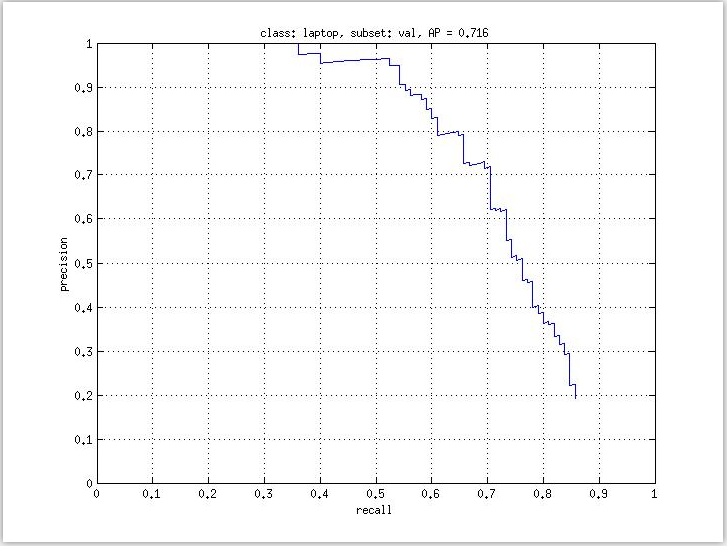


图20 AP曲线

4.2 基于ImageNet的训练集制作

在3.1小节中说到通用目标训练的流程的时候，说到制作训练集。制作训练集中除了在下载好的图像集中提取出部分图像外，还要根据提取出来的图像进行数据集分割。数据集分割就是把数据分为正负样本集，并将分割后的结果保存到txt文件中，以便在训练的时候读取数据。

下面是数据集分割的流程，以laptop为例，详见图21。

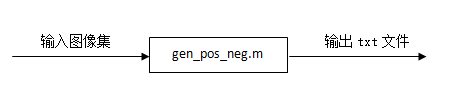


图21 数据集分割流程

（1）输入：在第三章中下载好的10个类别的图像集。

（2）过程：通过gen\_pos\_neg.m文件在laptop图像集中提取出200幅图像作为laptop正样本，其他9个类别的图像作为laptop负样本。

（3）输出：laptop\_train.txt，详见图22。

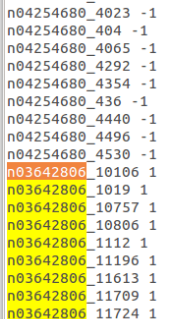


图22 laptop\_train.txt

在laptop\_train.txt中有两列数据，第一列是图片名称，第二列表示图片的正负性，1代表是正样本，-1代表是负样本。

4.3实验结果与分析

4.3.1训练结果分析

在经过一系列训练流程之后，选取的10个目标，由于样本选取原因，只训练出了9个目标模型，下面根据AP来分析一下9个模型的训练效果。详见表2，AP曲线图见附录C。

表2 各个目标模型AP值

|  |  |
| --- | --- |
| 目标类别 | AP值 |
| 飞机 | 0.816 |
| 热气球 | 0.660 |
| 船 | 0.757 |
| 蜡烛 | 0.400 |
| 汽车 | 0.800 |
| 牛 | 0.679 |
| 小刀 | 0.312 |
| 手提电脑 | 0.716 |
| 蜗牛 | 0.658 |

从AP值中可以看出，飞机、船、汽车的AP曲线是在0.75以上的，检测精度是较高的；热气球、牛、手提、蜗牛AP曲线是在0.6-0.7之间，检测精度比前者低；而小刀、蜡烛的AP曲线是低于0.4的，检测精度非常低。

从AP曲线可以知道，飞机、船、汽车在召回率小于0.8，精度是比较平稳的；在大于0.8之后精度就迅速下降；同时在召回率小于0.8时飞机和汽车精度都是从1.0开始缓缓下降，而船的精度是从0平稳上升的。热气球、牛、手提、蜗牛在召回率小于0.5之前精度都是比较平稳的，而在大于0.5-0.6之后，精度就开始迅速下降。小刀、蜡烛精度从一开始就下降，并没有太多的平稳期。

之所以会造成这样子的结果，原因可能有: 一、所选样本问题。所选图片可能具有光照差异、视点变化、遮挡问题、复杂背景等问题，提高了特征提取时的难度；除此之外，所选样本的不均匀分布也会造成训练出来的模型较差。二、算法原因。本文是基于DPM模型训练通用目标检测模型，可以看出，算法对于部件较多、机器类别的目标检测精度高，而对于生物、局部部件较少的目标物体检测精度明显的降低；三、验证集问题。模型训练完毕，进行验证的时候，由于验证集可能比较复杂带来AP精度较低。

4.3.2训练结果检验

上一小节中，根据AP曲线图大致分析了该算法对不同目标类别的检测精度，在这节中，将在生活或在网上找一些图片来检测，来验证下检测效果。详见图23 - 31。

飞机：检测精度高，目标包围盒覆盖率大。

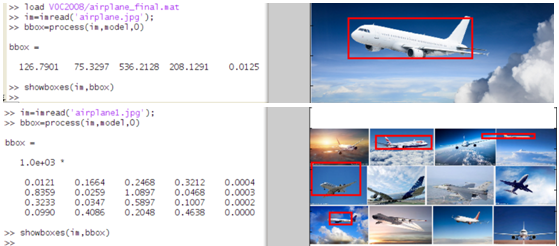


图23 飞机检测图

船：检测精度较高，只能检测出轮船。

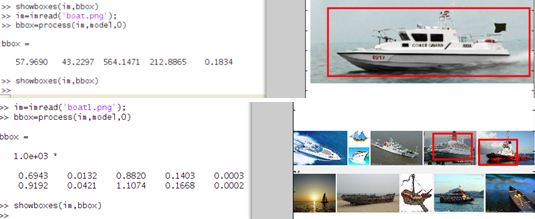


图24 船检测图

汽车：检测精度高，目标包围盒覆盖率较飞机的低。

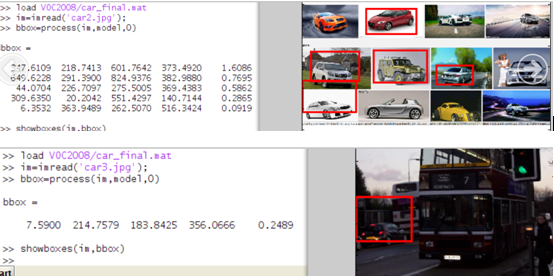


图25 汽车的检测图

热气球：检测精度高，目标包围盒覆盖率高。

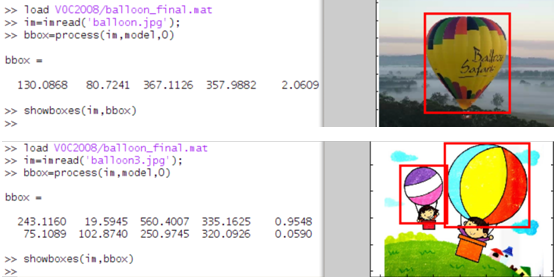


图26 热气球检测图

牛：检测精度高，无法检测出假的牛。



图27 牛检测图

手提电脑：检测精度高，覆盖率高，无法检测出合上的电脑或电脑背面。

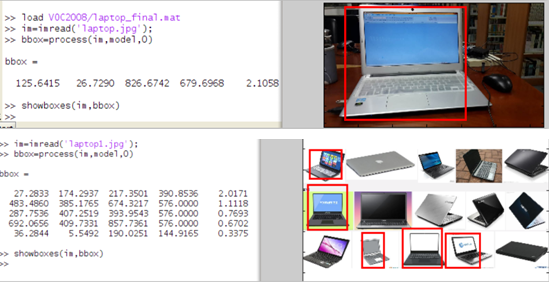


图28 手提电脑检测图

蜗牛：检测精度低，而且目标包围盒无法正确预测蜗牛整体，误将壳作为蜗牛。

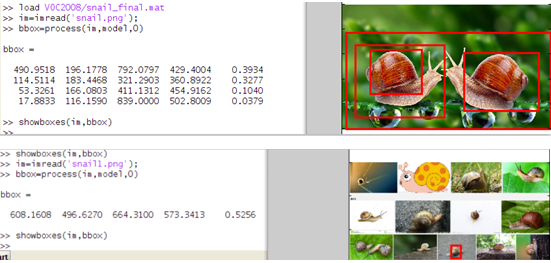


图29 蜗牛检测图

小刀：检测精度非常低，目标包围盒覆盖率低。

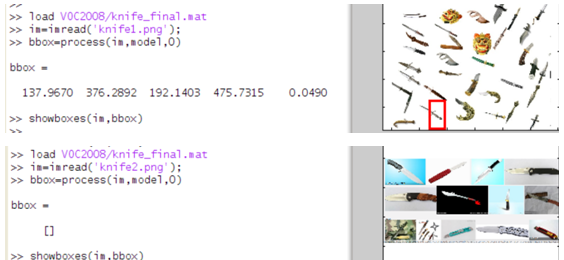


图30 小刀检测图

蜡烛：目标包围盒无法正确判断目标的所在。

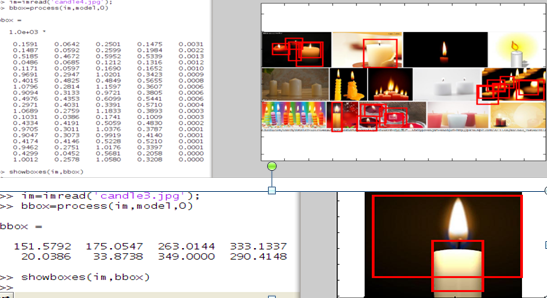


图31 蜡烛检测图

根据检测效果来看，主要有这几个问题：一、AP值较高的模型目标包围盒能很好的定位目标的位置，检测精度也高，但是对于各种形态的目标还不是能很好把握，如能检测到正面，无法检测到背面。原因可能是样本选择的时候不均匀或是样本数量不够。二、AP较低的模型，目标包围盒不能很好的覆盖目标的位置，而且容易对目标本身进行误判，比如只检测到目标物体的一部分或是根本不能识别图像中目标，检测精度也很低。

4.4本章小结

本章介绍了实验相关的一些情况，比如实验环境的搭建，数据集的分割，最后，对实验得到的模型进行分析和对日常生活中的物体进行检测。

5 总结与展望

5.1工作总结

本文基于对DPM了解后，基于ImageNet数据集训练了自己的通用目标检测模型，在训练自己的模型过程中加深了对DPM的深入了解，也对DPM算法在各类目标的检测率进行了验证。除此之外，在程序实现上，提出用Parfor并行计算替代for循环的串行计算，这将大大提高程序的运行速率。

现今对于特定的目标检测，如行人检测，或是目标结构比较简单的检测问题已经基本解决了，但是对于通用目标检测，在图像检测上，不管是多类别或是单类别多目标的检测都还是比较难，正在发展中的。本文通过DPM训练出的通用目标检测，从检测结果来看，AP值较高的模型，检测的限定框覆盖目标的程度较高，而对于AP值较低的模型，除了限定框的位置不能很好的定位目标之外，检测也会将目标的一部分当作是目标整体。而对于AP值较高的模型，也因为样本的原因，不能对目标各种形态（背面、侧面）都很好的检验出来，因此，前期样本的选取是非常重要的。

5.2研究展望

通用目标检测是现今一个比较有价值又对研究者具有挑战性的课题。到目前为止，已经有很多的通用目标检测算法，但是还没有一个还可以做到很好。这是因为不同类别的目标，结构不同，即使是同类别的目标，也会因为形态不同、光照强弱不同和背景不同而带来不确定性，这对算法提出了很高要求。如何在各种复杂多变的情形下保持稳定性是研究算法所需要解决的问题之一。本文的通用目标检测以基于可变形部件模型为基础，对通用目标进行模型训练。由于知识和精力有限，目前对算法只做到研究、分析阶段，还没有能力对算法进行改进，有如下几个方向亟待进一步的研究和发展改进。

（1）完备的训练样本。目前，识别率高的通用目标检测方法通常使用学习的方法，通过对各种目标样本的学习训练，得到一个判定模型，从而对待检测目标进行判定。所以，训练样本集的好坏关系到检测器的效果。好的训练样本集应该全面地包括要检测的目标和非检测目标样本，并且尽量减少样本的总数。因此，完备的训练样本集对于提高通用目标检测的精度是一个有效的方法。

（2）特征的准确提取。由于各种复杂多变的环境，会给目标图像带来多样性和可变性，如何能够有效地提取稳定的属于目标的特征是提高检测率的一个研究重点。例如，环绕在目标周围的场景，大家在研究时可能会作为干扰因素加以排除，但是，目标并不是孤立的存在的，场景也不是孤立的发生的，场景上下文能为目标检测带来很大的帮助，所以场景上下文的特征也能在目标检测发挥着非常重要的作用。

期待经过众多计算机视觉的研究人员的努力，通用目标检测能达到更高的检测率，更多地应用到多个领域。

参 考 文 献

何祥健,吴强,张淮峰.通用目标检测算法研究进展和评述[A].云南民族大学学报,2006:261-267

A Awan, R Dan, S Agarwal. Learning to detect objects in images via a sparse Part-base representation[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2004,26(11)：1475-1490.

A Leonardis,B Leibe,B Schiele.Combined Object Categorization and Segmentation with an Implicit Shape Model[C].ECCV Workshop on Statistical Learning in Computer Vision,2004:17-32.

A Mohan, C Papageorgiou, T Poggio. Example-based object detection in images by components[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2001,23(4): 349-361.

Andrew Zisserman, Christopher K.I.Williams, John Winn, Luc Vian Gool, Mark Everingham.The pascal visual object classes (voc) challenge 2010[J].International Journal of Computer Vision,2010,88(2):303-338.

Andrew Zisserman, Christopher K.I.Williams, John Winn, Luc Vian Gool, Mark Everingham. The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results[EB/OL]. http://host.robots.ox.ac.uk/pascal/VOC/voc2007/

B Schiele, C Wojek, P Dollar, P Perona. Pedestrian detection:A benchmark[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2009: 304-311

B Triggs, N Dalal. Histograms of Oriented Gradients for Human Detection[C].IEEE Computer Society Conference on Computer Vision and Pattern Recognition,2005:886-893.

CP Papageorgiou, M Oren, T Poggio. A General framework for object detection[C].Clinical Neurology and Neurosurgery, 2013, 108(6):573-5

David McAllester, Pedro F.Felzenszwalb.Object Detection Grammars[C].IEEE 13th International Conference on Computer Vision, 2011:691.

D Mcallester, D Ramanan, P Felzenszwalb. A discriminatively trained, multiscale, deformable part model[C].IEEE Conference on Computer Vision and Pattern 　　　　　Recognition,2008:1-8.

DW Kim, JK Hang, KK Sang. A recognition of vehicle license plate using a genetic algorithm based segmentation[C]. International Conference on Image Processing,1996,2:661-664.

ED Micheli, G Piccioli, M Campani. A robust method for road sign detection and recognition[J].Image and Vision Computing, 1996, 14(3): 209-223.

MA David, PF Felzenszwalb, RB Girshick, R Deva. Object detection with discriminatively trained part-based models [J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9):1627-1645.

N.Dalal.Inria person dataset[EB/OL]. http://pascal.inrialpes.fr/data/human/index.html

S Yan, TX Han, X Wang. An HOG-LBP human detector with partial occlusion handling[C]. IEEE International Conference on Computer Vision, 2009, 30(2):32-39.

# 附 录

附录A find\_obj.m

function find\_obj()

%find object id

objectlist = {'knife';'balloon';'snail'; 'candle'; 'soccer ball'; 'laptop'; 'airplane'; 'car'; 'boat'; 'cow'};

load('meta');

for i=1:10

oji=objectlist(i);

oji=oji{1};

for j=1:length(synsets);

idx=strfind(synsets(j).words,oji);

if length(idx)>0

pos=j;

fprintf([oji,' id is: %d\n'] , pos);

end

end

end

end

附录B download\_image\_imagenet.m

function download\_image\_imagenet()

%download image

load meta;

data=synsets;

path=data(ID).WNID;

synset\_url = ['http://www.image-net.org/api/text/imagenet.synset.geturls?wnid='data(ID).WNID];

[img\_url\_str,status]=urlread(synset\_url);

if status==1

img\_url\_array=regexp(img\_url\_str, '\r\n', 'split');

j=1;

k=1;

% path=['0000' num2str(k)];

while j<=200

if k<length(img\_url\_array);

[filestr,status]=urlwrite(img\_url\_array{k},['pictures/' path '\_' num2str(j) '.jpg' ]);

k=k+1;

if status==1

j=j+1;

end

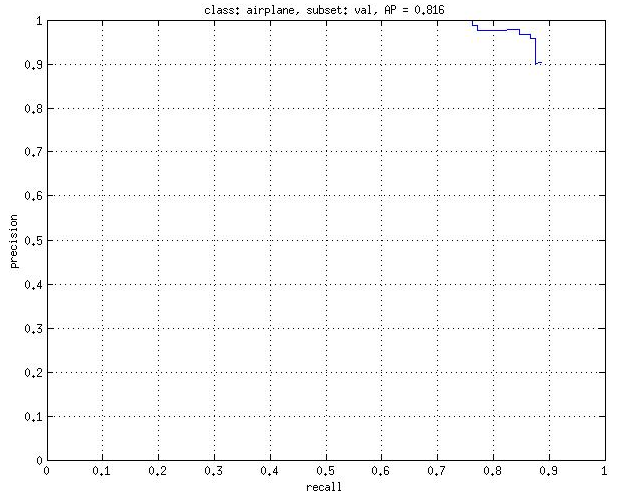
end

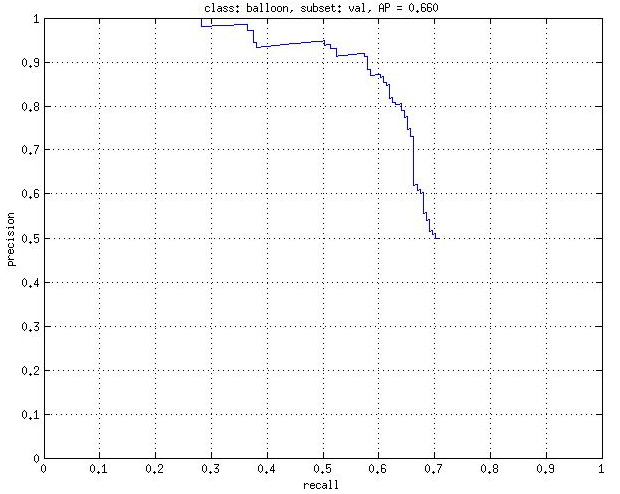
end

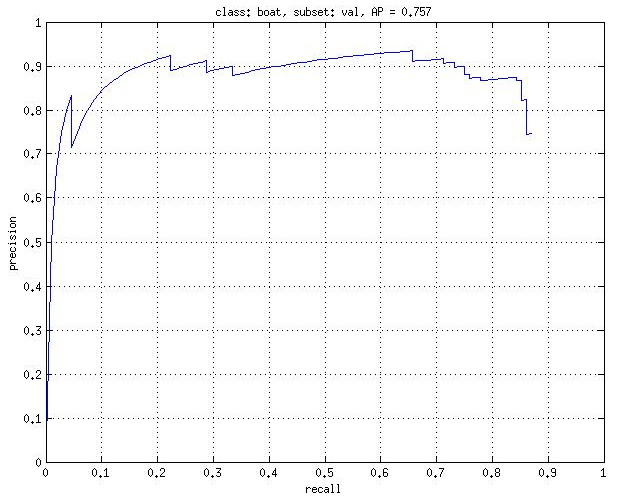
end

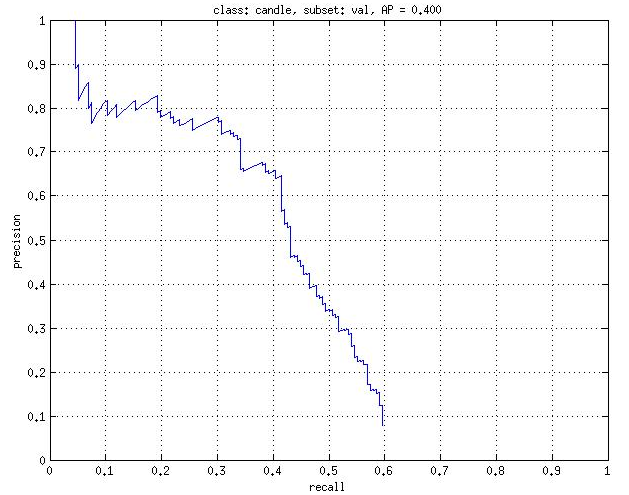
end

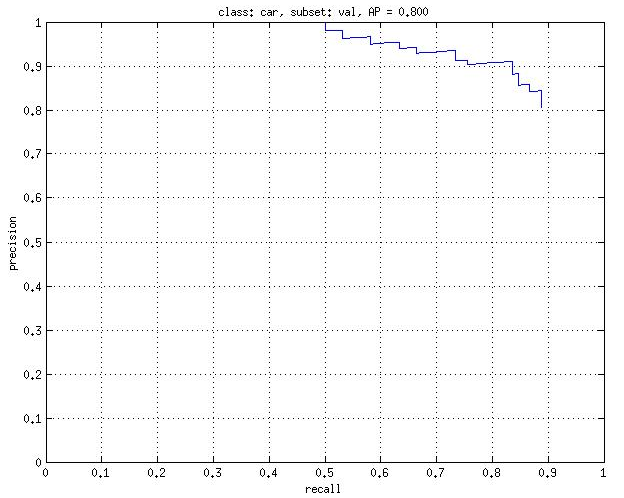
附录C 9个模型的AP曲线图

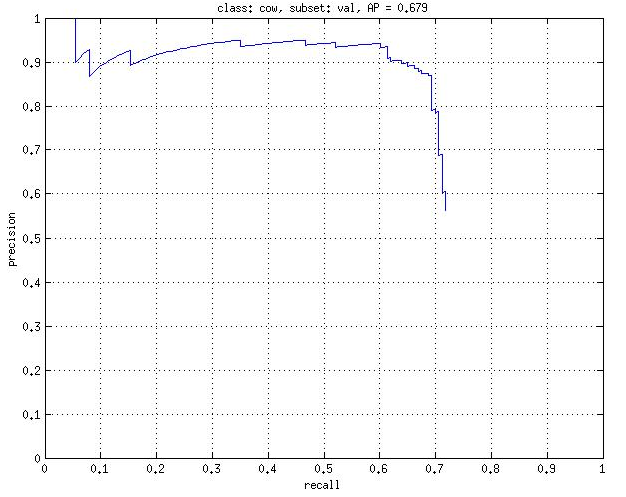


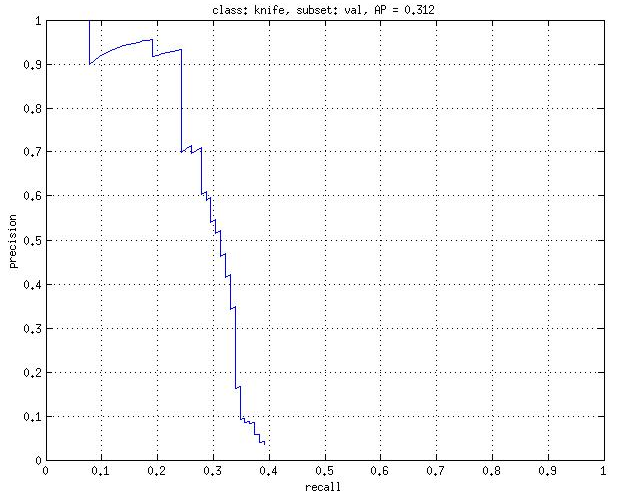


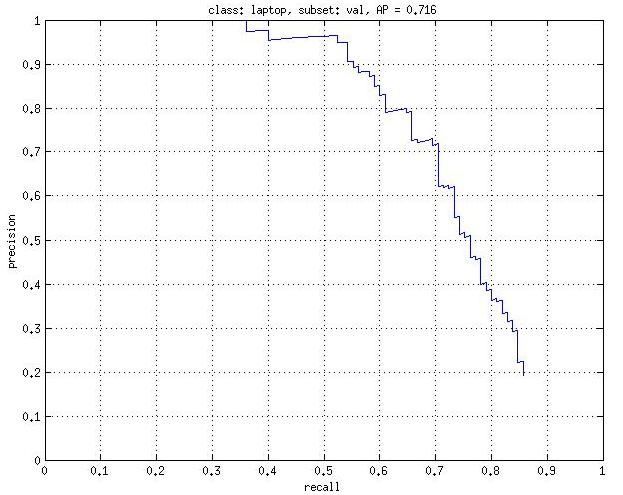


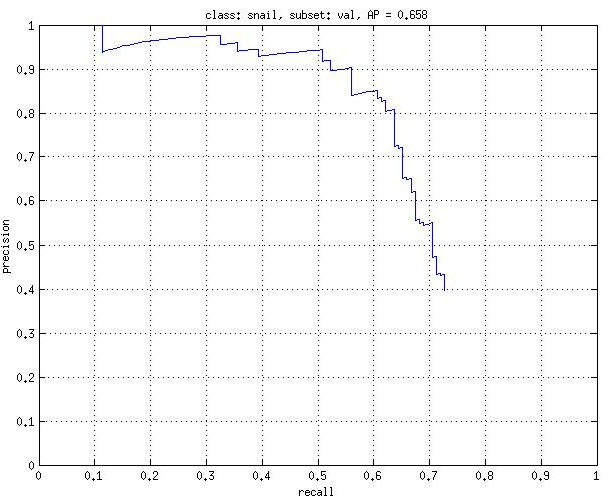












致 谢

时光飞逝，我的学生生涯也到达尾声了。在临近毕业之际，回顾过去4年的大学生活，有努力、有汗水、有失败也有收获。非常感谢华南农业大学提供了一个广阔的平台，锻炼、历练了我。华农这所大学的一种人文环境和奋斗精神深深地影响了我，“修德、博学、求实、创新”的校训我将铭记在心。

感谢我的指导老师刘老师。刘老师从论文的选题到撰写大过程中对我全程悉心指导。感谢刘老师给了我一个进入计算机视觉领域的机会，带领我做实验、指引我方向、教我不会的知识，给我提供了很多宝贵意见。

感谢我的家人。无论在什么时候都鼓励支持我，谢谢你们多年的养育。

最后，再次向所有关心、支持和帮助我的所有老师、同学、朋友、亲人表示感谢。