

学校编码: 10384

分类号 _____ 密级 _____

学号: 23020071151281

UDC _____

厦门大学

硕士学位论文

基于主动学习的图像协同标记算法研究

Image Co-labeling Based on Active Learning

刘礼锋

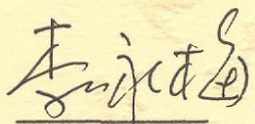
指导教师姓名: 曲延云 副教授

专业名称: 计算机应用技术

论文提交日期: 2010 年 5 月

论文答辩时间: 2010 年 月

学位授予日期: 2010 年 月

答辩委员会主席: 

评阅人: _____

2010 年 5 月

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

2010年6月5日

刘礼锋

摘要

图像的搜集与标记是计算机视觉研究的起点，一个标记好的图像集对分类器的学习及评估起着重要的作用。然而，现有的图像集及其图像标记基本上都是通过手工完成，费时耗力，不利于大规模图像集的建立。因此，研究图像自动标记关键技术对计算机视觉，模式识别等领域具有重要的科学价值。

本文研究了在已知少量标记样本的条件下，对大量没标记的图像进行标记的问题，提出了一种新颖的图像自动标记算法，该算法在协同学习的框架下，以主动学习的方式，采用两种特征互补的分类器对目标进行标记。主要的研究工作和创新点如下：

1. 提出对图像标记采用协同标记方式的算法。该算法采用 HOG 和 LBP 两种互补的特征形成两个分类器，对未标记图像进行协同标记，相对常用的单个分类器的标记算法，避免了错误的累积，提高了标记的精度。

2. 提出了在协同标记的基础上对学习模型进行更新的算法。该算法将增量学习的 SVM 应用于学习模型的更新，提高了分类器的性能。

3. 采用 GrabCut 算法对以方框标记的方式进行改进，以目标轮廓的方式对目标进行标记。

将本文算法应用于 Caltech 标准图像库及网上搜集的图像集，实验结果表明，通过主动学习的协同标记算法，能够较好的标记图像，标记精度优于单个学习模型的标记精度，学习模型的分类精度也随着学习样本的增加得到提高。

关键词：协同学习 增量支持向量机 图像标记 GrabCut

Abstract

Large databases of labeled images are the beginning of computer vision research; a labeled dataset plays an important role in learning and evaluating the classifier. However, the collection of image database and labeling images are basically done by hand currently, and it is trouble, labor-intensive and not conducive to establish a large image database. Therefore, the key technologies of image auto-labeling have intrinsic scientific value on computer vision and pattern recognition.

In this paper, we deal with the problem that how to label a large number of unlabeled images when knowing a small amount of labeled samples. We propose a new image auto-labeling algorithm which is under the framework of co-training and use active learning approach. We also use two features which are complementary to each other to train classifiers. The main researches and innovations are as follows:

1. We make use of co-label algorithm to label the image which uses two complementary features (the HOG feature and LBP feature) to train two classifiers, and then use these two classifiers to co-label the unlabeled images. Comparing with the labeling algorithm of a single classifier, our method avoids the accumulation of errors and improves the accuracy of labeling.

2. The proposed method updates our training model based on co-label. We use incremental SVM learning model to update and improve the classifier's performance.

3. We improve the label type using GrabCut algorithm which substitutes the rectangle label method with contour label method.

We test our algorithm on Caltech dataset and the images collected on the web. The experimental results show that our co-label algorithm based on active learning can label the image better, and has a great improvement in the label accuracy than a

single classifier model. In addition the accuracy of classification in our model can improve with the increasing of learning samples.

Keywords: co-training, incremental SVM, image labeling, GrabCut

厦门大学博硕士学位论文摘要库

目 录

第一章 绪论	1
1.1 研究背景	1
1.2 研究现状	3
1.2.1 流行的图像库介绍.....	3
1.2.2 图像标记相关工作介绍.....	7
1.3 本文的主要工作及内容安排	8
第二章 HOG+LBP 目标的特征表示	10
2.1 HOG	10
2.2 LBP	14
第三章 基于 SVM 的协同训练与增量学习	18
3.1 SVM.....	18
3.2 SVM 的增量学习.....	20
3.2.1 增量学习.....	20
3.2.2 增量 SVM.....	21
3.3 协同训练	27
第四章 基于主动学习的协同标记	30
4.1 初始训练	31
4.2 基于主动学习的协同标记算法详述	32
第五章 基于目标轮廓的图像标记	35
5.1 图切分算法	35
5.2 GrabCut 算法	37
5.2.1 GrabCut 数据模型.....	37
5.2.2 边界优化.....	40
5.2.3 迭代最小化.....	40
5.2.4 GrabCut 算法的实现步骤.....	40
5.2.5 GrabCut 算法的优点.....	41

5.3 基于目标轮廓的图像标记	42
第六章 实验结果与分析	44
6.1 图像库介绍	44
6.2 测试标准介绍	45
6.3 标记性能	47
6.4 分类性能	53
6.5 网络图像的收集和标记	54
6.6 GrabCut 轮廓提取	57
6.7 实验总结	59
第七章 全文总结及展望	61
7.1 总结	61
7.2 展望	62
参考文献	63
研究生期间参加的科研活动及科研成果	69
致谢	70

Contents

Chapter 1 Introduction	1
1.1 Research Backgrounds	1
1.2 Research status	3
1.2.1 Current DataSets	3
1.2.2 Research status of key technologies	7
1.3 Contributions and Outline	8
Chapter 2 HOG+LBP Representation	10
2.1 HOG	10
2.2 LBP	14
Chapter 3 Co-training and incremental SVM	18
3.1 SVM	18
3.2 Incremental SVM	20
3.2.1 Incremental Learning	20
3.2.2 Incremental SVM	21
3.3 Co-Training	27
Chapter 4 Co-active-learning	30
4.1 Initial training	31
4.2 Details of Co-active-learning algorithm	32
Chapter 5 Labeling Based on Coutour	35
5.1 Gragh Cut Algorithm	35
5.2 GrabCut Algorithm	37
5.2.1 Colour data modelling	37
5.2.2 Border Matting	40
5.2.3 Iterative energy minimization	40
5.2.4 Implementation steps of GrabCut	40
5.2.5 Advantages of GrabCut	41

5.3 Labeling Based on Coutour.....	42
Chapter 6 Experiments and analysis	44
6.1 Testing DataSet	44
6.2 Evaluation criteria	45
6.3 label ability	47
6.4 classify ability	53
6.5 Testing the performance on the web	54
6.6 Extrace contour using GrabCut	57
6.7 The analysis of results.....	59
Chapter 7 Conclusions and future work	61
7.1 Summary of this work	61
7.2 Future work.....	62
References.....	63
Publications	69
Acknowledgement.....	70

第一章 绪论

1.1 研究背景

图像库的建设对计算机视觉的研究起着重要的作用,它包括两方面的内容,一是图像的搜集,二是对图像中的目标进行标记。而图像库的作用主要表现在以下两个方面:一是作为训练集来训练出一个良好的分类器;二是作为测试集来评价分类器或算法的优劣。而一个好的图像库在这两个方面都扮演着重要的角色。常用的标准图像库有: Caltech-101^[1], Caltech-256^[2], PASCAL^[3], LabelMe^[4], LotusHill^[5], UIUC^[6]等,这些图像库的建设都是通过人工完成的,耗时耗力,成本高。特别是对大图像库的收集及标记,需要大量的人力及物力。而计算机视觉和图像处理的发展,使得计算机自动完成这个功能成为可能,所以图像自动标记技术最近成为了计算机视觉中的一个研究热点。

随着数码设备的普及,如拍照手机,数码相机, DV 等设备的平民化,使得人们更能随时随地地获取自己感兴趣的图像,也由于存储介质也越来越廉价,人们将会面对庞大的图像资源。同时随着因特网的发展,图像共享站点(如 Flickr, Facebook 等)开始出现,并且已变得越来越流行,这使得网络上所包含的图像资源也与日俱增,形成海量的图像资源,但是这些图像只是原始的、粗糙的、没有经过任何处理的图像,也没有预先经过分类,无法有效利用这些资源进行进一步的分析、搜索。本文的目的就是对图像进行自动标记,为图像的搜集整理奠定基础。

图像自动标记技术面临很多困难,迄今没有完善的解决方案,这些困难包括:

- 1、多种因素使同一物体在不同图像中存在很大的差异。比如亮度变化,如图 1.1(a);视角变化,如图 1.1(b);目标变形,如图 1.1(c);尺度变化,如图 1.1(d);目标部分遮挡,如图 1.1(e);背景嘈杂,如图 1.1(f)等等。同类物体之间存在较大差异,这要求识别模型既能体现同类物体之间的共性,又不能混淆相似的物体类别。

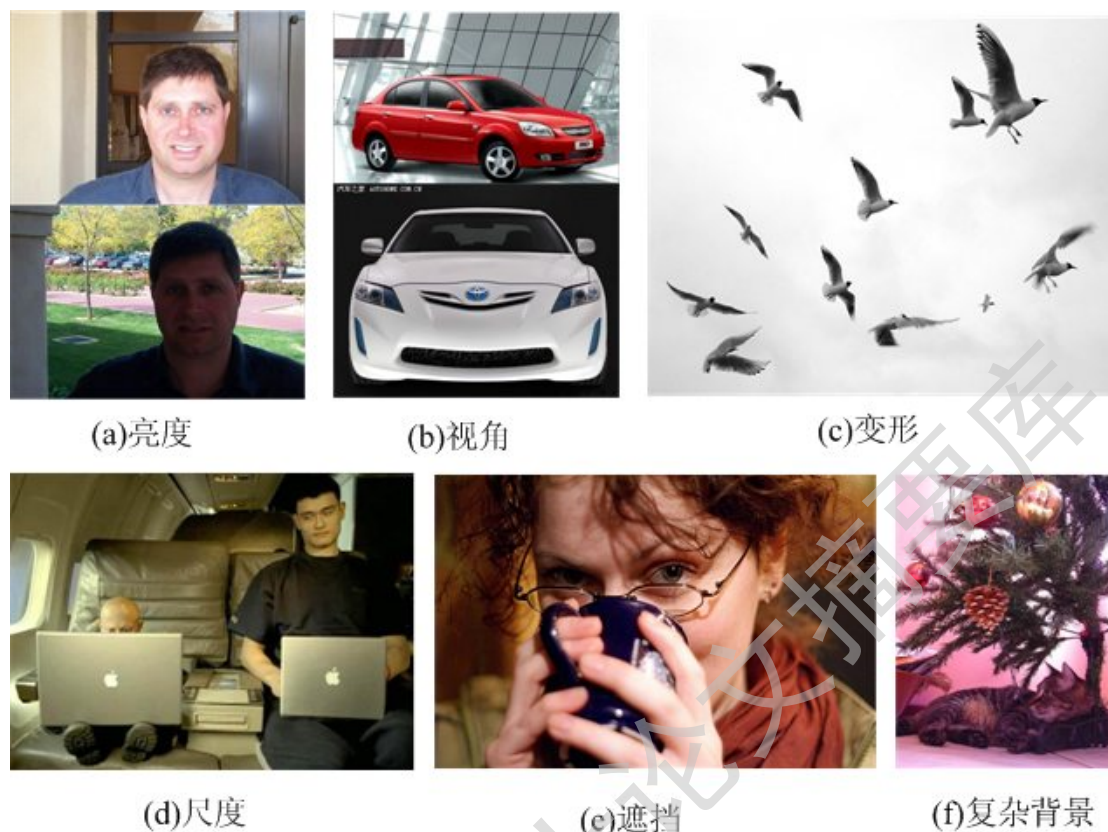


图 1.1 同类物体之间存在较大差异图像举例

2、标记的准确性。由于图像收集是用来训练或测试的，所以物体位置或轮廓标记的准确性是非常重要的，如果标记位置有所偏差或标记的大小不同，都会严重地影响学习模型的性能，然而这个标记的准确性又是难以保证的。

3、多类识别的困难。随着分类或识别算法的改进，其能处理的类别越来越多，由于训练和测试的图像库要求越来越全面，种类需求越来越多样化，每类图像的数量也越来越大，因此大量的类别增加了系统实现的难度。

4、时间复杂度和空间复杂度。因为图像是二维的信号，一幅图像所含像素数是成千上万的，有时需要对这些像素进行穷举的搜索，另外图像处理和分析常需要大量的运算，非常消耗时间和空间资源，所以处理效率问题也是图像自动标记技术的挑战之一。

5、机器学习本身也面临着很多挑战。图像自动标记技术通常都离不开学习算法，但是目前机器学习技术面临着诸多问题，比如学习能力、解决大规模图像库学习问题等等。

一个好的图像库应尽量多的包含目标的可能状态，对目标的标记应尽量准

确，这对于后期的目标检测、分类和识别具有重要的意义，直接影响着学习模型的精度以及评估。另外，随着互联网的发展，以及数字摄像设备的普及，导致海量图像数据的产生，图像标记的研究对图像分类，检索，标注等将产生深远的影响。

1.2 研究现状

1.2.1 流行的图像库介绍

随着科学技术的不断发展，正是由于图像库在计算机视觉方面的重要性，越来越多的人也致力于的图像库的收集工作，现在也有了一些比较好的人工收集的图像库，如 Caltech-101，Caltech-256，PASCAL，LabelMe，LotusHill 等，下面就对这些图像库分别作一些介绍：

1、Caltech-101

这是著名的用于物体分类和识别的图像库，由 Feifei Li, Marco Andreeto 及 MarcAurelio Ranzato 于 2003 年 9 月整理^{[1][7]}。其中有效物体共 101 类，再加上背景，共 102 类。每类中有大约 40 至 800 张图片，大部分图片有大约 50 张，每张图片的大小约为 300×200 。Caltech-101 所选取的 101 类物体包括了自然/人工，动物/植物，符号/实景等，充分体现了类与类的差异性，Caltech-101 是目前常用的标准图像库之一，很多目标识别算法利用该库检验其性能。在这个数据集上目前最好的物体识别结果由 UIUC 所提出的 LLC^[8]给出。最近又有更新的 Caltech-256^{[2][8]}问世，目标类别更多，图像数据相对更复杂。然而一个不足是类内变化范围不够显著，有人工痕迹。同类的图片往往同一尺度，基本上是单一目标图像，目标至少占图像大小的 30%，是标记后的图像，且大多数物体位于图片的中心附近，物体所受的变形与光照及拍摄的视角也大致相同，并且大部分图片都不是复杂场景(clutter scene)。其中的一些例子如图 1.2 所示。

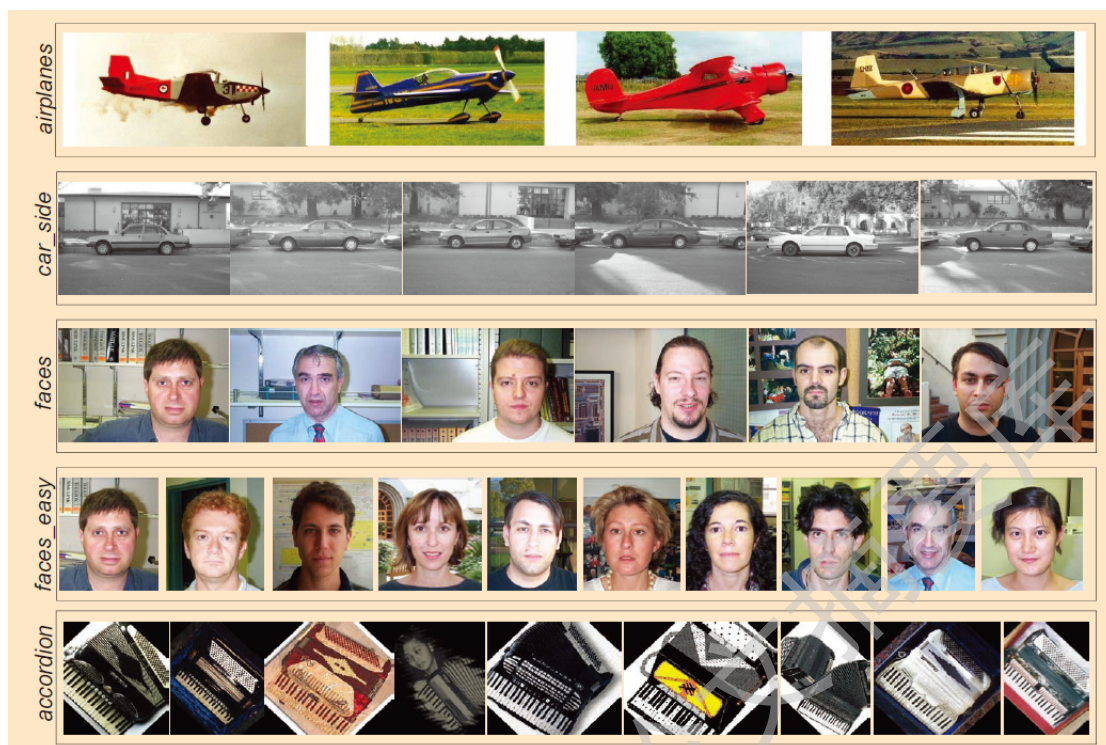


图 1.2 Caltech-101 图像库的一些例子

2、PASCAL (Visual object Class Challenge, VOC)

PASCAL 与前面的数据集不同，除了提供图像与标记，它更直接的目的在于收集了来自多个其他图像库的数据来进行竞赛，每年推出一个用于视觉对象分类挑战的训练和测试集合 (Visual object Class Challenge, VOC)，如 VOC2005^[3], VOC2006^[10], VOC2007^[11], VOC2008^[12], VOC2009^[13], VOC2010^[14]，供研究人员使用。同时，其还给出了测试的标准^[15]，如针对分类的 ROC 曲线，针对检测的 P/R 图及 DET 曲线图。与 Caltech 相比，PASCAL 图像数据的复杂性远远超出了 Caltech，同类图像包含了目标出现的多种状态，比如视角变化，光照变化，尺度变化，遮挡等，另外，同一幅图像上有多个目标。虽然只有 20 类图像，但每类图像集中所包含的图像数量很大，包含了目标尽可能多的状态。这更符合自然图像的性质，当然同时也加大了识别的难度。在竞赛中，参赛者不仅要为物体分类，而且要为物体定位，这使得该竞赛更充满了挑战性。其中的一些例子如图 1.3 所示。

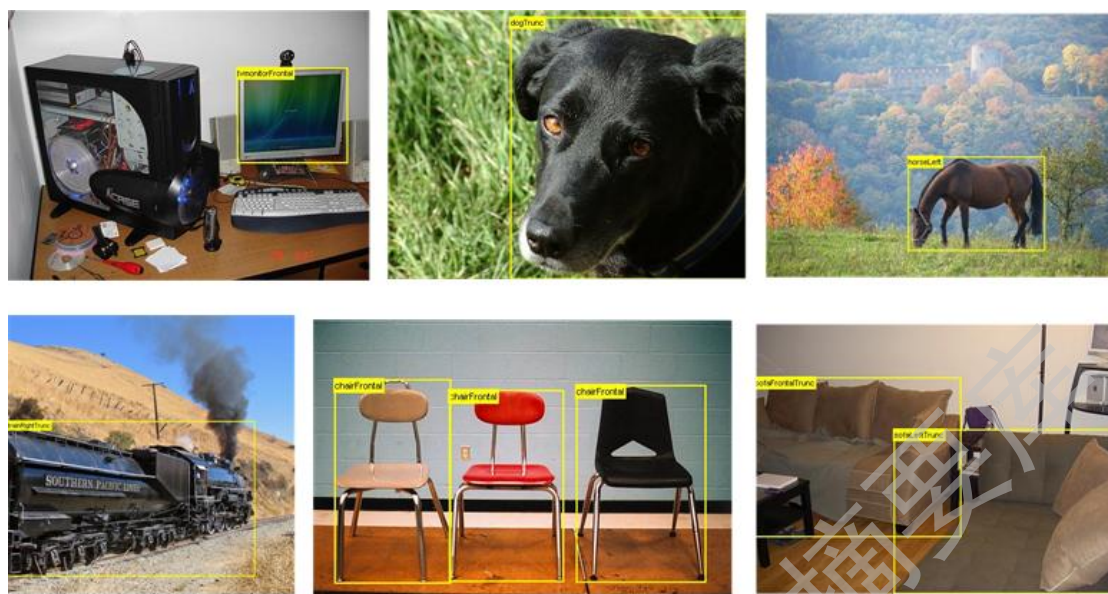


图 1.3 PASCAL 图像库的一些例子

3、LabelMe

LabelMe 是麻省理工学院 CSAIL 实验室建设的图像库^{[4][16]}，利用互联网搜集图像、标记图像。它包括了自然图像、各种场景以及具有不同视角不同姿态的物体数据。相对 Caltech-101 来说，这个数据集的变化远远多于 Caltech-101，能更好的作为通用图像物体检索算法的一个基准。它包含 106,739 张图像，其中 41,724 幅已经被标记，所有图像中被标记的物体总数达到 203,363 个。但是标记结果只有用多边形描述的粗糙简单的外轮廓数据。该数据集利用互联网由网页访问者以参与游戏的形式进行标记，虽然这样可以收集大量的各种场景，包括各种目标的图像，但由于上传者的评判标准不一，且带有很多主观因素，则也带来了标记质量的控制与管理等问题。因此很难把它作为一个通用的图像物体数据集。还有，在这个数据集中，某些物体类含有非常多的图像，而某些类的图像数目则非常之少，都需要耗费很大的人力来收集和选出图像。其中的一些例子如图 1.4 所示。

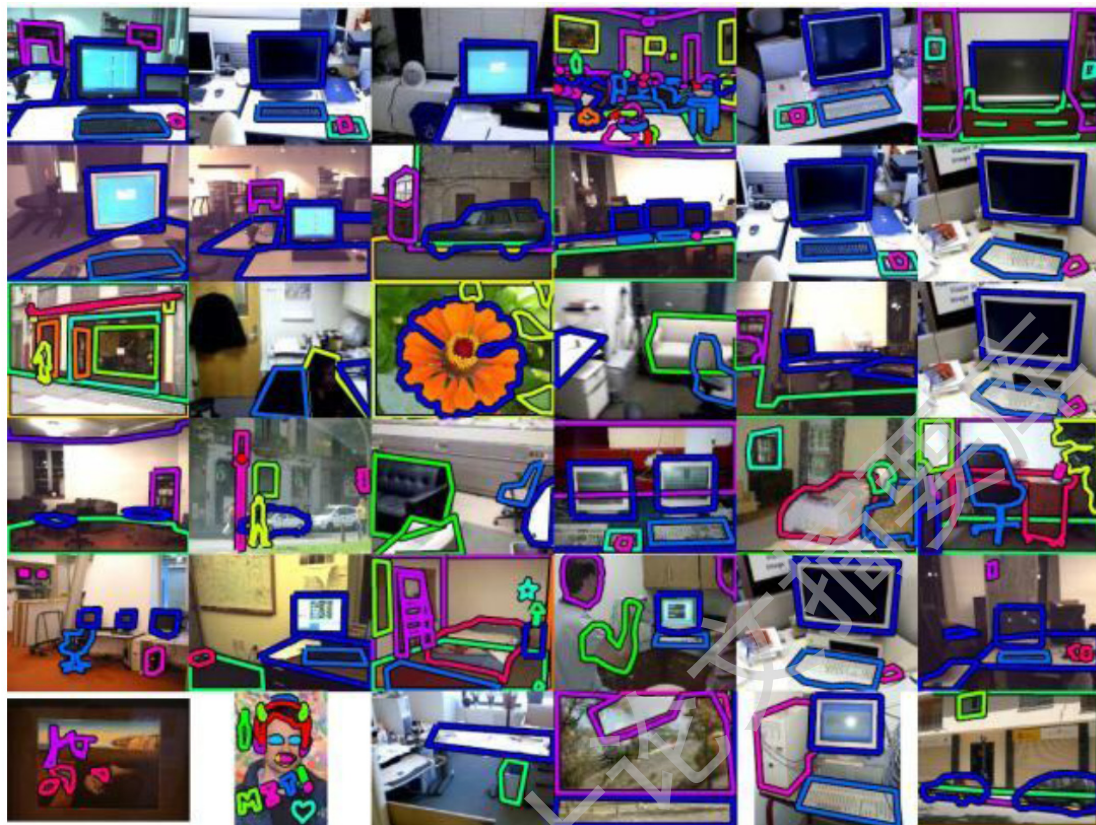


图 1.4 LabelMe 图像库的一些例子

4、LotusHill (LHI Image Database for the ground truth of vision)

这个图像库是由中国莲花山计算机视觉和信息科学研究院建设的^{[5][17]}。主要研究目的是建立手工标注图像数据库的理论、方法和规范，目标是建立一个世界上最大的百万幅人工标注图像数据库，为目前计算机视觉问题提供丰富的实验原料和科学的验证平台。目前图像库中已经收录了几万幅不同种类，不同尺度，不同视角，不同场景的真实图像。截止到 2007 年 2 月 2 日为止，共收集了 636, 748 张图像，3, 927, 130 个物体，包含场景、航空图片、物体、人脸、文字、视频等。但是也是由人工进行收集并人工进行标记，特别是对这种特别大的图像库，更是非常的耗时耗力。其中的一些例子如图 1.5 所示。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库