

学校编码: 10384
学 号: 31520141153305

分类号____密级____
UDC____

厦 门 大 学

硕 士 学 位 论 文

基于特定类别空间约束的弱监督目标检测

Category-Aware Spatial Constraint for Weakly Supervised
Object Detection

沈云航

指导教师姓名: 纪荣嵘 教授

专 业 名 称: 计算机技术

论文提交日期: 2017 年 5 月

论文答辩时间: 2017 年 5 月

学位授予日期: 2017 年 月

答辩委员会主席: _____

评阅人: _____

2017 年 5 月

厦门大学博硕士学位论文摘要库

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学博硕士学位论文摘要库

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）：

年 月 日

厦门大学博硕士学位论文摘要库

摘要

目标检测 (object detection) 是计算机视觉领域里一个非常重要的研究问题。随着近年来深度卷积神经网络的发展, 其中基于深度学习的目标检测算法在性能上取得了巨大的进步。但是目前最先进的目标检测算法需要带有精确目标物体位置标签的数据来训练模型, 而这种标签信息需要花费大量人力物力来标注, 同时也会引入人工标注偏差。

本文的研究内容是基于弱监督学习的目标检测问题, 即没有精确的目标物体位置标签, 只用图像的分类标签来学习目标检测器。基于弱监督学习的目标检测有着广泛的应用和重要的意义, 也是近年来的计算机视觉领域的热门研究点。当前基于弱监督学习的目标检测算法大多是基于局部、候选区域层次的信息。对此, 本文提出一种算法探索和结合无监督的目标物体全局的形状和位置信息来协助模型的训练。本文的主要研究内容可以概括为以下三点:

1. 本文提出特定类别的像素梯度图。在训练过程中, 本文提取图像的特定类别的像素梯度图。基于特定类别的像素梯度图, 模型可以粗略估算目标物体的形状和位置;
2. 本文利用目标物体的粗略估计和候选区域位置的关系, 提出了候选区域的空间约束。基于候选区域的空间约束, 模型能把特定类别的全局信息和候选区域的局部信息引入模型的学习过程中;
3. 本文提出一种多中心正则化来惩罚预测得分比较高的候选区域的特征和对应类别的特征中心的不一致。多中心正则化使得模型训练更加的稳定。

本文的算法没有提高网络模型的复杂度, 也没有使用额外的监督信息。最后, 大量的实验结果表明本文的方法取得了优异的弱监督目标检测和定位性能, 并超过目前所有最先进的方法。

关键词: 弱监督学习; 目标检测; 深度学习

厦门大学博硕士学位论文摘要库

Abstract

Visual object detection is sitting on the core of computer vision research. Recently with the development of deep convolutional neural networks, object detections based on deep learning have achieved great progresses. However, current the state-of-the-art object detection algorithms require training dataset with object-level label to learn the models. It is time-consuming and labor-intensive to obtain such object-level label and can introduce some annotator bias.

This paper focus on problem of object detection based on weakly supervised learning, which is lack of object-level label and only have image-level in dataset for training the object detector. Weakly supervised object detection has wide application and important significance, and has seen a surge of attention from the computer vision community in the recent years. Current weakly supervised object detection algorithms are based on local, instance-level information. Therefore, we propose to exploit and incorporate unsupervised object global shape and location information to assist the model training. The follows are the main contents of this paper:

1. We propose a novel category-specific pixel gradient map. During training, we extract category-specific pixel gradient map, based on which we roughly obtain location and shape of target object;
2. We combine the rough estimation of object and location of regions to propose the spatial constraint, based on which we can introduce category-specific global information and region local information to the learning process;
3. We propose a multi-center regularization to penalize violations between category centers and high-score regions, which makes the

learning more stable.

Our algorithm neither arises the complexity of model, nor requires addition supervised information to learning. Finally, extensive experiments show that our approach achieves outstanding performance of object detection and localization, and significantly outperforms the state-of-the-art methods.

Key words: Weakly Supervised Learning; Object Detection; Deep Learning

厦门大学博硕士学位论文摘要库

目录

摘要.....	7
Abstract.....	9
第一章 绪论	15
1.1 课题研究背景与意义.....	15
1.2 论文的研究目标和内容.....	17
1.3 论文的组织结构.....	19
第二章 研究概述	21
2.1 目标检测的国内外研究现状.....	21
2.2 多示例学习的国内外研究现状.....	26
2.3 弱监督学习的国内外研究现状.....	28
2.4 基于弱监督学习的目标检测的国内外研究现状.....	30
第三章 基于特定类别空间约束的弱监督目标检测算法	37
3.1 网络结构.....	38
3.2 特定类别的像素梯度图.....	42
3.3 候选区域空间约束.....	45
3.4 多中心正则化.....	51
3.5 实验验证和分析.....	53
3.5.1 数据集和评估标准.....	53
3.5.2 实验细节.....	53

3.5.3 训练伪代码	55
3.5.4 定量评估	56
第四章 总结与展望	65
4.1 本文工作总结	65
4.3 今后的研究方向	65
参考文献	67
附录 攻读硕士学位期间发表论文及科研情况	73
致谢	75

Contents

Abstract (Chinese)	7
Abstract (English)	9
Chapter 1 Introduction	15
1.1 Background and Significance of the Research	15
1.2 Objective and Content of the Thesis	17
1.3 Structure of the Thesis	19
Chapter 2 Literature Review	21
2.1 Object Detection Review	21
2.2 Multiple-Instance Learning Review	26
2.3 Weakly Supervised Learning Review	28
2.4 Weakly Supervised Object Detection Review	30
Chapter 3 Category-Aware Spatial Constraint for Weakly Supervised Object Detection	37
3.1 Network Architecture	38
3.2 Category-Specific Pixel Gradient Map	42
3.3 Region Proposal Spatial Constraint	45
3.4 Multi-Center Regularization	51
3.5 Experiment And Analysis	53
3.5.1 Datasets and Evaluation Protocol	53
3.5.2 Implementation Details	53

3.5.3 Pseudo Training Code.....	55
3.5.4 Quantitative Evaluations	56
Chapter 4 Conclusions and Future Work.....	65
4.1 Conclusions	65
4.2 Future Work	65
References.....	67
Appendix: List of Publications.....	73
Acknowledgements.....	75

厦门大学博硕士学位论文摘要

第一章 绪论

1.1 课题研究背景与意义

目标检测是计算机视觉领域中的一个基础性的研究课题，主要需要解决图像里有“什么”和在“哪里”的问题。近十年来计算机视觉里的目标检测问题得到极大的研究发展，大量基于深度学习的新方法被提出。在现实世界中，不同类别的物体的视觉差异可能是很小的，而同一类别的不同物体的差异不仅受物体物理属性变化的影响，还受成像条件变化的影响。例如，花在生物学上是十分多样的，不同个体间的形状，颜色和纹理等属性是千变万化的。在现实场景中，目标物体往往伴随出现视觉上类似的背景形势，或者目标物体只占据整个场景的很小部分，或者目标物体可能被其它物体遮挡，这些各种可能的情景对目标检测任务构成很大的挑战。我们知道目标检测可以分为两个关键的子任务：目标分类和目标定位。前者回答了图像里有“什么”的问题，后者回答了目标在“哪里”的问题。目标分类任务负责判断图像中是否存在特定类别的目标物体，输出一系列带分数的标签表示特定类别的目标物体出现在图像里的可能性。目标定位任务负责搜索图像中特定类别的目标物体的位置和大小。目标检测有着广泛的实际应用，比如智能视频监控，增强现实、基于内容的图像检索和机器人导航等等。除此之外，目标检测也是很多高级计算机视觉任务的重要前提，比如：身份识别和验证、场景分析和理解等等。综上所述，目标检测无论是在计算机视觉领域里还是在实际应用中，都具有非常重要的意义。因此在最近的二十年里，众多科研人员密切关注目标检测问题并投入大量的精力对其进行研究。而伴随着强大的深度学习和强劲的硬件平台发展，近十年来和目标检测相关的课题和研究不仅有增无减，而且模式多样化，每年都有最新的研究成果发表，最新的实际应用公布。尽管如此，目前目标检测算法的性能(检测准确率和检测速度)跟人类相比起来还是相差非常远。所以说，目标检测问题没有被完美的解决，依旧是计算机视觉领域里一个重要的、具有挑战性的一个研究课题。

通常的目标检测算法是基于有监督学习。近年来大多数目标检测相关的研究

也是关注基于有监督学习的目标检测算法。除此之外另外一个值得关注的方向是基于弱监督学习的目标检测的研究。弱监督学习（Weakly Supervised Learning）是机器学习领域和模式识别领域里一个非常热门研究方向。事实上，根据训练数据里监督信息的精细度，本文大致把机器学习划分为三种：有监督学习、弱监督学习以及无监督学习。根据训练数据里监督信息的形式，弱监督学习其实又可以细分为多示例学习和半监督学习。多示例学习的数据是由若干个只有类别标签的包构成，而每个包包括了若干个没有任何标签的示例。假如一个包里所有的示例至少有一个是正样本，那么这个包的标签是正的。假如一个包里所有的示例都是负样本，那么这个包的标签就是负的。另一方面，半监督学习则是使用少量有监督信息的数据和大量无监督信息的数据一起进行学习的问题。虽然多示例学习和半监督学习是有很大差别的，但是它们都是只需要部分或者不完整的监督信息来进行学习。我们可以看出，弱监督学习是处于有监督学习和无监督学习两个极端的中间。实际上在现实生活中，通常带有弱监督信息的数据是远远多于有监督信息的数据。因此基于弱监督的学习的算法有着广泛和重要的应用场景。但是目前对于基于弱监督学习的目标检测的研究工作还是比较少的。而且基于弱监督学习的目标检测算法的性能也差强人意，因此基于弱监督学习的目标检测是一个十分值得研究的课题。

通常训练目标检测需要大量人工标注的精细监督信息：目标类别标签和目标位置标签。目标类别标签通常用只包含 0 和 1 的向量来表示，1 代表图中存在对应的目标，0 表示图中不存在对应的目标。而目标位置标签通常用方形包围盒的形式来表示。通常只需要四个坐标就可以确定一个包围盒。这种精细的目标位置标签通常需要付出大量的人力物力来获取。在人工标注包围盒的过程还会引入标注偏差进而影响训练结果。事实上，只带有目标类别标签的数据是比较容易获得或者标注的，比如用户在网络上传图像，通常会对图像添加标题或者描述。我们可以从互联网获得大量的弱监督标签信息的数据。因此，一个自然的想法就是只使用只有目标类别标签的数据来训练目标检测器，这也正是本文要研究的问题。

当前基于弱监督学习的目标检测仍存在着严峻的挑战。总体来说，基于弱监督学习的目标检测带来的挑战性主要是以下两个方面：鲁棒性和计算复杂性。

类内表观差异和类间表观差异是影响基于弱监督学习的目标检测的鲁棒性

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