

分类号: TP

单位代码: 10335

密 级: 无

学 号: 21625004

浙江大学

硕士学位论文



中文论文题目: 基于深度学习的机械臂分拣系统
及其云平台研发

英文论文题目: Robotic Arm Sorting System Based on
Deep Learning And It's Cloud Platform

申请人姓名: 朱雨贺

指导教师: 陈子辰教授

合作导师: 傅建中教授、姚鑫骅副教授

专业名称: 机械制造及其自动化

研究方向: 智能制造技术及装备

所在学院: 机械工程学院

论文提交日期 2019 年 1 月 25 日

基于深度学习的机械臂分拣系统 及其云平台研发



论文作者签名:

签字

指导教师签名:

签字

论文评阅人 1: XXX XX XXXX

评阅人 2: X X XX XXXXX

评阅人 3: XXX XX XXXXXX

评阅人 4: XXX XX XXXXX

评阅人 5: XXX XX XXXXX

答辩委员会主席: 唐三藏 功佛 洛阳大慈恩寺

委员 1: 惠 能 方丈 曹溪宝林寺

委员 2: 智 顓 方丈 天台山国清寺

委员 3: 法 藏 大和尚 洛阳佛授记寺


委员 4: 道 济 和尚 临安灵隐寺

委员 5: 降 龙 尊者 天竺大雷音寺

答辩日期: 2019 年 3 月 15 日

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得浙江大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。


学位论文作者签名：


签字日期： 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解浙江大学 有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权浙江大学可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名：

导师签名：

签字日期： 年 月 日

签字日期： 年 月 日

致 谢

两年半的求学生涯，在各位老师和朋友的帮助和支持下，即将画上一个圆满的句号。回顾过往，虽然辛苦却也收获良多，转眼已到毕业的时候，心中难免不舍。在此，我要向所有支持我的人表达衷心的感谢。

首先，我要感谢我的导师陈子辰教授对我的培养。陈老师渊博的知识、认真的科研精神、诲人不倦的高尚品德以及谦和的人格魅力对我产生了深远的影响。陈老师虽然忙身于教学与科研重任，但仍然经常指导我，特别注重我们独立思考能力以及工程实践能力的培养。此外，陈老师还关心我们的生活，注重我们人格的培养。有幸能成为陈老师的学生，我不仅学会了如何科研，更学会了如何做人。

我还要特别感谢傅建中教授。傅老师在我的整个硕士生涯中，对我的科研生活起到了极大的知道作用。傅老师身为实验室的带头人，为实验室的师生营造了一个自由向上，积极进取的科研氛围。在傅老师的带领下，实验室科研成果不断。也得益于实验室自由的环境，在科研之外，各个同学在各个领域都有所建树，全面发展。同时，傅老师在我的毕业设计中，细心指导了我的毕设选题、创新点发掘、论文撰写，使我能够顺利完成毕业论文的撰写。生活中，傅老师平易近人，亦师亦友。

我还要特别感谢姚鑫骅副教授。姚老师在我的硕士生涯中影响巨大。姚老师在我的日常课业、科研项目以及挑战杯项目中起到了很大的指导作用。姚老师具有渊博的知识、严谨的治学态度，是我在科研项目中的带头人。同时，姚老师不仅在科研中给予了我莫大的帮助，在生活以及心理方面，姚老师也不断地鼓励着我。帮助我顺利完成了硕士学业。

感谢沈洪垚副教授、贺永教授、吴森洋老师、郑亚平老师、周芳丽老师，在课题组进行科研的过程中，他们在很多方面都影响和帮助着我。

同时，我还要感谢所有教导、帮助、关心过我的老师。在整个本科和硕士学习生涯中，他们给予了我莫大的帮助，他们的帮助也让我成长了许多。

感谢实验室已经毕业的王博、延健磊、代晓萌、饶成晨、兰刘健、刁怀东等师兄师姐在科研、求职以及生活中对我的帮助。感谢实验室的林志伟、栾丛丛、孙扬帆、牛小淼、张德明、刘加朋、张承谦、谢明君、孙元等博士生给予我的帮助。感谢与我同一届的孙伟俊、

刘森鑫、黎清雨、干胜、赵耀、刘丞哲、夏能、王鹏、严梦玲、黄天玺、周瑞剑等同学，我们共同求学于同一个实验室，互相帮助，共同克服科研路上的难题。感谢史璇珂、吕健冉、杜旺哲、王润秋、刘诗怡、刘冰、林炜奕、方泽华、叶潇翔、田佳陇、李霄铨等师弟师妹，在给予我帮助的同时，也为实验室带来了新的活力。

感谢女朋友申佳佳的帮助、包容与鼓励，我们一起求学、求职，她在我低谷时给予我温暖的鼓励，在我迷茫时为我出谋划策，是她让我更加自信、更加坚定自己的方向。

感谢同寝室的武欣、黎清雨、苟华伟同学，感谢他们在生活上给予我的帮助和支持。

最后，我要感谢的是我的家人，感谢他们这么多年来对我的养育和培养。二十年寒窗苦读，是他们的支持和鼓励让我走到了今天。他们默默付出，不图回报，他们的爱是最宝贵的财富。

如今，我即将离开学校，步入下一段新的旅程，衷心感谢这些年来所有帮助过我的人，谢谢！

朱雨贺

2019 年 1 月于求是园

摘 要

待填

关键词：深度学习，目标检测，分拣系统，机械臂，云平台

Abstract

The quick brown fox jump over the lazy dog.

T_EX

Keywords: T_EX

目录

致谢	I
摘要	III
Abstract.....	V
目录	
第 1 章 绪论	1
1.1 论文研究背景及意义	1
1.1.1 研究背景	1
1.1.2 研究意义	2
1.2 国内外研究现状	3
1.2.1 目标检测算法研究现状	3
1.2.2 基于视觉的自动分拣系统研究现状	5
1.2.3 深度学习云平台研究现状	7
1.3 论文研究内容与架构	8
1.3.1 论文研究内容	8
1.3.2 论文架构	10
1.4 本章小结	11
第 2 章 基于深度学习的机械臂分拣系统设计	13
2.1 系统整体架构	13
2.2 系统硬件选型	14
2.2.1 相机	15
2.2.2 目标检测模型训练用服务器	15
2.2.3 目标检测运行硬件平台	15
2.2.4 机械臂选型	15
2.2.5 系统硬件架构	15
2.3 系统模块间通信机制	15

2.3.1	USB 通信	15
2.3.2	ROS 通信	15
2.3.3	串口通信	15
2.4	目标检测算法选型	15
2.5	相机标定及机械臂控制模块设计与实现	15
第 3 章	图像处理模块设计与实现	17
3.1	图像算法理论	17
3.1.1	感知算法	17
3.1.2	YOLOv3	17
3.2	图像采集和数据获取	17
3.3	数据标注	17
3.4	目标检测模型训练与评估	17
3.5	目标检测模型部署	17
第 4 章	基于深度学习目标检测模型的云平台设计与实现	19
4.1	深度学习云平台方案设计	19
4.1.1	云平台整体架构	19
4.1.2	基于 Http 协议的通信方法分析	19
4.1.3	云平台性能需求分析	19
4.2	基于 Flask 的 Web 服务端设计与实现	19
4.2.1	云平台服务端架构	19
4.2.2	基于 Linux 的 Shell 脚本实现云平台模型训练	19
4.3	基于 Web 前端的客户端系统设计	19
4.3.1	客户端网页需求分析	19
4.3.2	客户端页面分析、设计与实现	19
4.3.3	Web 前端与服务器端的通信实现	19
第 5 章	实验研究	21
5.1	机械臂分拣系统性能实验	21
5.1.1	分拣延时实验	21
5.1.2	分拣准确率实验	21
5.2	基于深度学习的目标检测模型云平台性能测试	21
5.2.1	Web 前端性能测试实验	21
5.2.2	服务端性能测试实验	21

第 6 章 总结与展望	23
6.1 总结	23
6.2 展望	23
参考文献	25
附录 A.....	27
附录 B.....	29

第 1 章 绪论

1.1 论文研究背景及意义

1.1.1 研究背景

人工智能的迅速发展将深刻改变人类社会生活，改变世界。为此，国务院在 2017 年 7 月 8 日向各省、自治区、直辖市人民政府等印发了《新一代人工智能发展规划》^[1]。此规划是为了推动人工智能快速发展，造福民生。其中，“人工智能+”是其中的重要环节。所谓“人工智能+”，就是将人工智能与各个传统行业深度融合，深入改造各个传统行业，帮助传统行业转型，创造新的产业形态。而智能制造则是人工智能与制造业的深度融合。十八大以来，“制造业转型升级”等战略被逐步提出，但对于中国制造来说，要实现转型升级，首先要考虑由劳动密集型向技术密集型转变。而人工智能技术与制造业的融合可以代替大量劳动力，实现工业自动化，助力“中国制造 2025”，帮助中国制造转型升级。

工业机器人，按照 ISO 8373^[2] 的定义，是面向工业领域的多关节机械手或多自由度的机器人。工业机器人通过预先编写的程序可以实现重复性动作，以此可来代替一些重复性的人工工作，是制造业中非常重要的劳动力代替工具。根据国际机器人联合会发布的 2012 年世界机器人研究报告，在 2011 年年底，世界上运行中的工业机器人数量达 1153000。可以说，工业机器人的每次技术进步，都能带来制造业工作效率质的提升。

分拣作业是制造业中非常常见的一种作业场景，由于其重复性强、作业场景单一的任务特性，分拣作业成为工业机器人的重要应用场景之一。传统的工业分拣采用人工方法，不但耗时耗力，而且无法满足自动化长时间作业，影响生产效率的提升。而传统的工业机器人自动分拣系统，采用预先编程的工业机器人进行工业分拣，虽然能够实现重复性动作的复现，但由于工业机器人无法根据实际情况改变自身动作，所以这种系统对分拣件的摆放位置有严格的设定，且该系统只能进行分拣，无法进行分类。因此，近年来，越来越多的学者将机器视觉应用到工业机器人当中，使机器人具备人眼功能，能够根据工件位置自适应地调整机器人动作，并且能够实现分拣和分类。基于机器视觉的工业机器人系统具备更好的鲁棒性，因此越来越多的机器视觉工业机器人投入到实际应用中，不仅提高了生产产

品质量，更保证了工业化生产的效率^[3]。

机器视觉技术是利用机器代替人眼做各种测量和判断的技术^[4]。工业自动分拣系统主要由机器人、视觉传感装置、图像采集装置和图像处理系统构成^[5]，其中机器人模块主要执行图像处理系统所返回的指令，其他模块构成了机器视觉模块，负责获取所需识别物体的位置及类别。由于机器人模块受环境影响较小，而图像处理模块需要满足工业环境中的实时性和快速性要求，而且需要对工厂复杂多变的环境有较高的适应性。因此，机器视觉中的目标检测算法是整个视觉自动分拣系统中的核心。

人工智能的高速发展必定伴随着对各行各业的改造。以深度学习为代表的算法也在不断颠覆各种传统机器视觉算法。将深度学习应用到基于视觉的工业分拣系统中，不仅能够提高视觉识别的效果，而且得益于深度学习的端到端的特性，系统的可移植性也可以得到提高。本文研发的基于深度学习的机械臂分拣系统及其云平台，将深度学习应用到自动分拣系统的核心模块：目标检测，并且研发出一套基于深度学习，用于训练目标检测算法的云平台，降低视觉应用门槛，具备更强的通用性。

1.1.2 研究意义

随着“中国制造 2025”的提出，中国正迈向“制造大国”到“制造强国”的道路。中国制造业自动化转型迫在眉睫。工业机器人，作为制造业中不可或缺的生产工具，在制造业自动化中扮演着重要角色。工业机器人任何技术上的提升必定带来制造业生产效率的提升。因此，对工业机器人的技术研究对实现“中国制造 2025”具有重要意义。

作为工业机器人应用的重点和难点，工业分拣系统具备大多数工业机器人应用的技术应用场景。工业自动分拣系统的研究突破势必带动工业机器人技术的进步。工业自动分拣系统主要包括图像采集模块、图像处理模块、机器人模块。其中，图像采集模块主要用于获取分拣件图像信息，作为图像处理模块的输入；图像处理模块主要根据图像采集模块采集得到的图像信息，进行计算，得到分拣件位置和类别信息，并将该信息发送给机器人模块执行；机器人模块根据分拣件位置和类别信息，对分拣件进行抓取并放到合适位置。这其中，图像处理模块是整个系统的中枢，决定了整个系统的实时性和准确性。一个好的图像处理模块带来的系统性能提升是其他模块无法带来的。

分拣系统中用到的图像处理算法主要为目标检测算法。目标检测算法可以根据是否应用深度学习技术，将之分为传统目标检测算法和基于深度学习的目标检测算法。传统的目标检测算法需要人为构造特征，且针对不同检测物体需要构造相应的特征。当需要检测的物体数量变多并且经常发生变化时，传统的目标检测方法就会显得捉襟见肘。而基于深度

学习的目标检测算法可以实现端到端的训练和预测。同时，深度学习端到端的特性既方便了模型的训练，又方便模型的移植。针对不同的检测场景，只需要标注相应数量的数据，利用迁移学习^[6]技术，在 ImageNet^[7]的预训练参数上进行 fine-tune，即可得到良好的可用模型。

鉴于以上特性，根据当前工业自动分拣系统及我国制造业特性，提出了本论文的基本设想，即开发一套基于深度学习的机械臂分拣系统，同时将深度学习模型的训练过程开放到云平台。本论文的设想分两部分：一是将基于深度学习的目标检测算法应用到工业分拣系统当中，验证其可行性及效果；二是开发基于深度学习的目标检测模型训练云平台，帮助制造业从业者生成他们所需的目标检测模型。相对应地，本论文有如下两部分意义：一是研究深度学习在工业分拣系统中的应用效果，提高工业自动分拣系统的分拣质量及鲁棒性；二是从我国制造业从业者深度学习知识普及度不高的事实出发，将深度学习模型的训练过程封装为对制造业从业者友好的网页界面，从而普及深度学习在工业自动分拣系统甚至是整个制造业中的应用。

1.2 国内外研究现状

1.2.1 目标检测算法研究现状

目标检测是机器视觉中的一个重要问题，在轨迹跟踪、自动驾驶、工业分拣等生活及工业领域都具有重要的应用和研究价值。早期的目标检测算法主要是人工构造特征，通过模板和传统机器学习方法进行检测。然而，2012年，Hinton等人提出了 AlexNet^[8]，在 ImageNet 的表现超出第二名 41%，将深度学习的革命性效果带到大众眼前。深度学习在图像分类任务上的良好表现，促使各界学者纷纷在各自的领域应用深度学习，由此催生出一批基于深度学习的优秀算法。目标检测任务也不例外。

自目标检测的概念提出以来，国内外学者针对目标检测算法进行了一系列探索。在深度学习兴起之前，传统的目标检测算法大致分为目标实例检测和传统目标类别检测两类^[9]：

1. 目标实例检测，即通过模板和人工提取的特征点，获得模板与检测对象的对应关系，检测出目标对象。
2. 目标类别检测则通过使用传统的机器学习方法，根据人工提取的特征和选定的分类器，检测出实现定义好的有限几种类别。

2004年，Lowe等提出在目标实例检测方面应用极为广泛的 SIFT^[10]算法，该算法使用高斯模糊来实现尺度空间，使用高斯差分函数进行极值检测，再通过对各个因素的判定和

筛选,得出匹配度高、抗噪性强的关键点。但该算法存在复杂度高、速度慢、对于特征不明显图像很难提取特征点等问题。针对这些问题,Ke 等人同年提出了 PCA-SIFT^[11] 算法。该算法引入 PCA 方法对 SIFT 算法中的向量进行降维,以提高匹配效率。但由于降维损失了部分信息,因此该算法虽然提高了匹配效率,但匹配效果具有局限性。同样基于 SIFT,2006 年提出的 SURF^[12] 算法则引入了 Hessian 矩阵,减小了计算量,提高了目标检测的速度。而在目标类别检测方面,使用较多的则是基于 AdaBoost^[13] 的系列算法。AdaBoost 算法是一种 Boosting 算法,能够将多个弱学习器,通过调整训练集样本权重的方法,组合为一个强学习器。

传统的目标检测算法,其目的都是在人工提取丰富特征点的前提下,尽可能减少计算量,从而提高计算效率,提高识别速度。但人工提取特征虽然易于理解,简单直观,但无法应对大量类别的识别,在目标识别体发生变化时,需要针对性地再次进行繁杂的特征设计和提取工作。而基于深度学习的目标检测算法,使用神经网络提取图像底层和高层特征,不仅能够提取出更加丰富、表达性更强的特征,而且不需要人工参与特征的提取,还能做到端到端的训练和预测。

基于深度学习的目标检测算法分为两类。一类是基于分类的目标检测算法(two-stage),另一类则是以回归的方式来进行目标检测的算法(single-stage)。顾名思义,two-stage 的方法,采用两步来进行目标检测:1. 生成可能区域(Region Proposal)并且使用 CNN 提取特征。2. 使用分类器分类并修正位置。这类方法的代表是 R-CNN^[14]。2014 年, Ross 等人提出了 R-CNN。该算法使用 Selective Search 获得候选区域,然后进行归一化,作为 CNN 网络的输入。再使用神经网络获取候选区域的特征,最后利用多个 SVM 分类器进行分类。R-CNN 大幅度提高了目标检测的准确率,但由于 R-CNN 是对所有的候选区域进行特征提取和计算,多个候选区域之间的重叠区域重复计算了多次,导致了大量重复计算,降低了运算效率,并且由于计算过程复杂,计算过程中需要存储大量中间数据,需要消耗大量的存储资源。因此,R-CNN 的实时性不强,并且非常占用存储资源。针对 R-CNN 的重复计算性问题,2014 年,何凯明等人提出了 SPP-Net^[15]。SPP-Net 不同于 R-CNN 对所有候选区域进行特征提取,而是使用 CNN 对整张图片进行一次特征提取,大大减小了计算量,并且在 R-CNN 的最后一个 CNN 层之后,加入了 SPP 层。但 SPP-Net 依然没有解决 R-CNN 占用过多存储空间的问题。2015 年, Fast R-CNN^[16] 在 SPP-Net 的基础上进行了改进,将 SPP 层进行了简化,将分类和边框回归问题进行了合并,并且引入 SVD 分解,减小了计算量。Fast R-CNN 相对于 SPP-Net 提高了计算速度,减小了内存占用,但仍然使用 Selective Search 的方法选取候选区域,并进行大量计算,其计算速度依然不尽如人意。针对候选区域选择的问题, Faster R-CNN^[17] 使用 RPN 网络代替了 Selective Search,标志着基于深度学

习的目标检测算法走上了真正端到端的计算。但由于 Faster R-CNN 仍然使用了 Fast R-CNN 中的 ROI 层,导致其在小目标的检测任务上效果不尽如人意。并且由于 two-stage 本身思路的限制, two-stage 系列算法中最快的 Faster R-CNN 也只能达到 5FPS。因此,研究者们提出了另外一种思路: single-stage。即将目标检测和分类问题直接转化为回归问题,去掉选择候选区域及提取特征这一步,直接通过整图得到检测物位置和类别。这一系列算法的代表是 YOLO^[18] 算法。YOLO 算法将整张图片划分为 $S \times S$ 个网格,各个网格负责检测中心落在该网格内的检测物体,包括检测物位置和类别信息。该算法通过舍弃了一定的精确度换取了速度的大幅度提升,其检测速度最高可达 45FPS。但检测精度比 Faster R-CNN 要差。同年提出的 SSD^[19] 算法,则结合了 YOLO 和 Faster R-CNN 的优点,在保证检测精度的同时,兼顾检测速度。YOLO 算法发表的次年,即 2017 年,YOLO 的作者又提出了 YOLOv2^[20] 算法,该算法通过在卷积层后面加入 BN 层、加入 K-Means 聚类方法等方式,同时提高了检测精度和速度。该算法代表了 2017 年业界最先进的目标检测算法。2018 年,YOLOv3^[21] 基于 YOLOv2 做了一些设计细节的改进,在保持高检测速度的前提下,再次提高了检测精度。总之,针对不同的应用场景, two-stage 算法和 single-stage 算法各有其适用范围。没有不好的算法,只有合适的使用场景。

1.2.2 基于视觉的自动分拣系统研究现状

基于视觉的分拣系统,利用机器视觉技术,对目标进行识别与跟踪,将目标位置和类别信息反馈给机械臂控制模块,从而控制机械臂完成目标工件的分拣工作。基于视觉的分拣系统在国内外的研究已经比较成熟,应用场景正在不断扩大。国内相关研究起步较晚,但近年来也在加大研究力度,逐步缩小与国外的研究差距。

由于自动分拣系统在各个领域应用十分广泛,而不同领域所需使用的算法及机械臂均不相同,因此,在各个领域,均有相应的自动分拣系统研究成果。如水果分拣领域,该文献^[22]提出了一种基于视觉的水果自动分拣方案,能够根据分类器分出水果的质量好坏,根据水果横径尺寸讲水果分为大、中、小果,同时能够根据水果表面颜色信息对水果成熟度做出判断。在制药领域,有文献^[23]设计出一种药片自动分拣系统,通过药片图像的预处理,进行药片缺陷的检测与分类,并在此基础上实现了药片的自动分拣。制造业领域,自动分拣系统主要用于工件的分拣,并且在分拣过程中可以做到工件的无损检测。此文献^[24]设计了一个工件智能分拣平台,能够对工件进行在线动态分拣,使得分拣效率大幅度提升。

国内的自动分拣系统研究中,对其中的目标检测算法也做了许多创新性的研究。比如,伍锡如等人^[25]在进行工件目标检测时,结合了传统图像算法和深度学习算法。在工件定位



(a) FANUC M-1iA 型号机器人

(b) FlexPicker 机器人



(c) 物流自动分拣机器人

图 1-1 自动分拣机器人实例

阶段使用传统图像算法做匹配以确定工件位置，在工件分类时则使用 CNN 进行分类。此外，由于我国电子商务告诉发展，国内每年上亿包裹，也催生了快递领域自动分拣系统的应用。如该文^[26]中提到的物流分拣机器人，通过视觉技术，可根据商品的重量及发往地等信息进行快速分拣，极大地提高了快递发货周期，提高服务水平。其中运用到了大量的视觉技术，如数量检测、形状识别等。

目前，国内有多家提供自动分拣解决方案的公司。如深圳辰视智能科技有限公司，该公司采用传统机器学习方法与深度学习结合的方式定位图像中的目标，引导机械臂对目标采取相应的操作，以高效低成本的方式为机器人厂家提供机器视觉相关的解决方案。

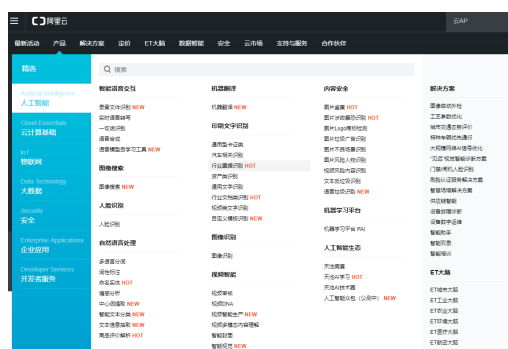
国外方面，著名的机器人生产厂商如日本的 FANUC、EPSON，瑞典的 ABB 以及德国的 KUKA 等企业均有自己的自动分拣机器人产品^[27]。FANUC 公司适用于分拣作业的机器人型号为 M-1iA 型，如图 1-1(a) 所示。该型号机器人具有重量轻、结构紧凑的特点，可以安装在各种复杂的环境中。具有快速的响应速度，可以快速调整工件姿态。该机器人配备配套的软件包 iRVision 内置有各种视觉库，可以完成目标检测功能，该机器人配备该软件包可实现基于视觉的自动分拣功能。EPSON 公司生产的机器人种类繁多，具有代表性的为 S5L 型号机器人，其机械结构刚性、功能强大、运行平稳。基于 Vision Guide 视觉软件可以实现视觉功能。ABB 的 FlexPicker 机器人则擅长拾取操作，运行速度快，抓取精度高，非常适用于工件分拣，如图 1-1(b)所示。同样，该机器人需要适配 Cognex 公司生产的配套视

觉软件方能实现自动分拣功能。

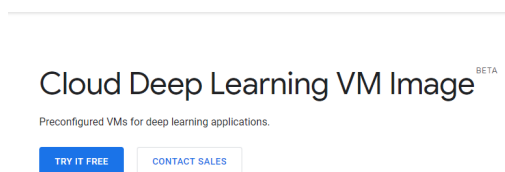
1.2.3 深度学习云平台研究现状

随着深度学习的高速发展，人工智能逐渐深入到我们的生活当中。而深度学习的高计算资源需求意味着深度学习的高准入门槛，因此，为了满足广大开发者的深度学习计算需求，一批深度学习云计算平台应运而生。

国内的代表是云计算龙头：阿里云^[28]。其界面如图1-2(a)所示，阿里云主要为国内中小企业提供云服务。具体来说，就是将阿里巨量的服务器资源开放出来，以出租的方式租赁给企业。租赁服务主要包括存储、数据库、企业应用、开发者API等。其中深度学习服务主要以API的形式对外开放。如果用户需要自行进行深度学习计算，则需要自行申请带有好计算资源的服务器，通过远程连接到服务器，通过服务器上自带的深度学习框架，自行编写代码进行训练。



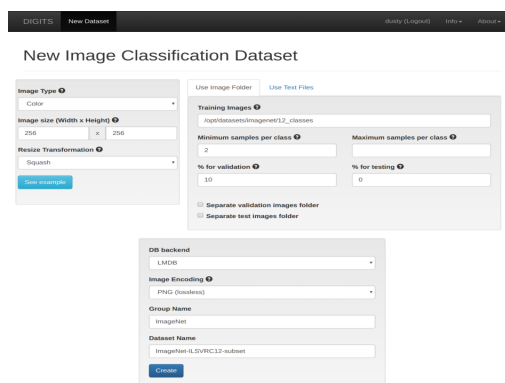
(a) 阿里云界面



(b) 谷歌云界面



(c) 亚马逊云界面



(d) 英伟达云界面

图 1-2 国内外云平台实例

国外云计算起步较早，如美国的谷歌、亚马逊、英伟达等公司均提供优秀的云服务。其中亚马逊云^[29]占据了美国市场最大的份额。亚马逊云界面如图1-2(c)所示。亚马逊云为客

户提供功能十分丰富的云服务，其中，深度学习相关资源的提供方式和阿里云类似。值得一提的是英伟达云（Nvidia GPU Cloud, NGC）^[30]，其界面如图1-2(d)，NGC 主要提供深度学习计算服务，并且与阿里云和亚马逊云只提供计算资源不同，NGC 将常用的深度学习算法训练过程进行了封装，通过网页端以选择框和文本框的形式供用户自定义深度学习训练参数配置。同时提供了常用的深度学习数据集供用户使用。用户只需要选择配置和数据集即可训练自己的深度学习模型，大大降低了深度学习的计算和知识门槛。

综上可知，国内外云计算平台的深度学习服务主要通过 API 的方式提供，深度学习定制化模型一般只面向具备开发能力的开发者。而我国国内制造业从业者普遍缺乏云计算平台的使用能力及深度学习知识，即使能够使用云计算提供的高计算资源，也很难去定制自身需要的深度学习模型。因此，针对国内制造业从业者缺乏深度学习知识而制造业又需要深度学习模型落地的具体情况，开发一款面向制造业的封装良好且易于上手的云平台十分必要。

1.3 论文研究内容与架构

1.3.1 论文研究内容

自动分拣系统主要由图像采集模块、图像处理模块和机器人模块构成，其中图像处理模块的核心为目标检测算法。传统的自动分拣系统，其主要采取传统的目标检测算法，针对特定的检测工件，人工构造特征，进行模板匹配工件位置，然后使用这些特征作为分类器的输入再进行分类。总而言之，传统的分拣系统，其采取的目标检测算法需要根据不同的工件人为构造特征，一旦工件发生变化，需要重新提取特征。这样的系统可移植性较差。随着深度学习的快速发展，基于神经网络的端到端的训练方式逐步改造着视觉方面的各类算法。神经网络可以在训练过程中自动提取底层和高层特征，能够很好的表征图像特征，无需人为提取和构造特征，且效果也优于传统算法。由于神经网络的特征提取特性，当工件类别发生变化时，无需耗费过多人力即可完成模型的移植，因此，基于深度学习的目标检测算法具有很好的可移植性。此外，基于国内制造业从业者缺乏深度学习知识和技能的现状，业内急需一款面向制造业的低门槛深度学习云计算平台。出于以上几点，本论文设计并实验了一个基于深度学习目标检测算法的机械臂分拣系统，并开发了一个用于定制深度学习目标检测模型的云平台。

本论文一共分为两部分。一是机械臂分拣系统的设计与实验；二是用于定制深度学习目标检测模型的云计算平台。

其中，基于深度学习的机械臂分拣系统需要实现如下功能或模块：

1. 图像获取模块，实时获取图像信息，并以一定的通信方式将图片实时传递给图像处理模块。
2. 图像处理模块，接收图像获取模块传递来的图像信息，通过训练好的目标检测模型预测该图像中工件的位置及类别，并以一定的通信方式传递给机械臂模块。
3. 机械臂执行模块，包括机械臂控制模块和机械臂；接收当前图像中的工件位置和类别信息，将其转化为相应的控制指令，控制机械臂将工件抓起，并放到该工件类别相对应的位置。

定制深度学习目标检测模型的云计算平台需要实现以下功能：

1. Web 端方面，提供文件上传功能，用于接收用户自行标注的数据集，以定制化目标检测模型。
2. 同样是 Web 端，以选择框或文本框的形式给出用户可自定义的超参数选项，方便用户定制化自己需要的模型特性。
3. 服务器端，接收 Web 端发送的各项参数，并根据参数执行相应的 shell 脚本，调用 GPU 训练深度学习目标检测模型，并且能够将训练过程中的各项参数实时同步到 Web 端，方便用户可视化模型的训练过程。

基于深度学习的目标检测模型虽然具有效果好、可移植性强等诸多优点，但其所需计算资源较大，难以在普通的嵌入式平台或台式机上运行，而使用 GPU 服务器进行实时计算会让整个系统显得笨重，因此，如何在嵌入式平台进行深度学习模型的预测运算是一个难点。其次，图像采集模块、图像处理模块、机械臂控制模块需要高效的通信机制，以便以高 FPS 完成采集、处理、执行的流程。云平台方面，则需要探究 Web 前端和服务器端的通信机制，及服务器端拿到配置信息后，如何调用 GPU 去训练深度学习模型。

针对以上问题，本论文针对以下方面进行了深入研究：

1. 基于深度学习的目标检测模型的训练。首先搭建图像采集模块，并对所要检测的工件进行图像采集。针对这些图像，使用特定的图像标注工具进行标注。选取 YOLOv3 作为目标检测模型，选取其开源框架 Darknet 进行训练，训练使用两块 GTX 1080Ti 显卡。并对其训练过程中的各项参数如 Loss、IOU 等值进行可视化分析。
2. 机械臂控制模块的设计实现。由于相机与机械臂所在空间的零点不一致，因此系统在使用之前需要进行标定。将机械臂和相机的坐标统一化。之后分析如何根据相机中工

件的位置信息计算出机械臂下工件的位置信息，然后研究机械臂的控制方法，控制机械臂执行相应的动作。

3. 机器人通信机制研究。使用机器人操作系统（Robot Operation System, ROS）实现图像采集模块、图像处理模块、机械臂控制模块的高效通信。并研究如何将 YOLOv3 模型封装为 ROS 节点，加入整个系统通信网络。
4. 云平台结构设计与实现。包括 Web 前端页面的设计与实现，文件上传功能的实现，使用 JavaScript 获取前端组件信息，并封装为 JSON 数据发送给服务端。服务端则需要接收并解析 JSON 配置数据，并编写相应的 Linux shell 脚本去训练定制化的深度学习模型，并将模型训练过程中的信息实时同步到 Web 前端。

1.3.2 论文架构

针对课题的研究内容，本论文各个章节的主要内容安排如下：

第 1 章，首先阐述了本论文的研究背景和研究意义；然后介绍了目标检测算法、基于视觉的自动分拣系统和深度学习云平台的国内外研究现状；最后给出了本论文的主要研究内容和论文架构。

第 2 章，主要从整体上阐述了本课题中基于深度学习的机械臂分拣系统的设计方案；并介绍了系统的硬件选型及依据，阐述了系统所使用的通信方案；之后根据本系统的具体条件，选择合适的深度学习目标检测算法；最后介绍了如何进行相机标定和机械臂控制。

第 3 章，重点介绍图像处理模块。首先介绍了系统中使用的图像处理算法理论；然后介绍如何使用图像采集模块获取数据集并进行标注；最后是深度学习目标检测模型的训练、评估和部署。

第 4 章，首先介绍深度学习云平台的整体设计方案和通信方法，并进行了云平台的用户需求分析；之后分别介绍了云平台服务器端和 Web 端的实现方案。

第 5 章，分别就本论文的两大部分进行实验。第一部分对自动分拣系统的延时和准确率进行了实验，并和采用传统目标检测算法进行了对比，确定深度学习方法的优越性。第二部分则对深度学习云平台服务端和 Web 端进行了稳定性测试实验。

第 6 章，对本论文所做的工作进行了总结，并对未来可作的改进工作进行了探讨。

论文的章节架构图如图 1-3 所示。

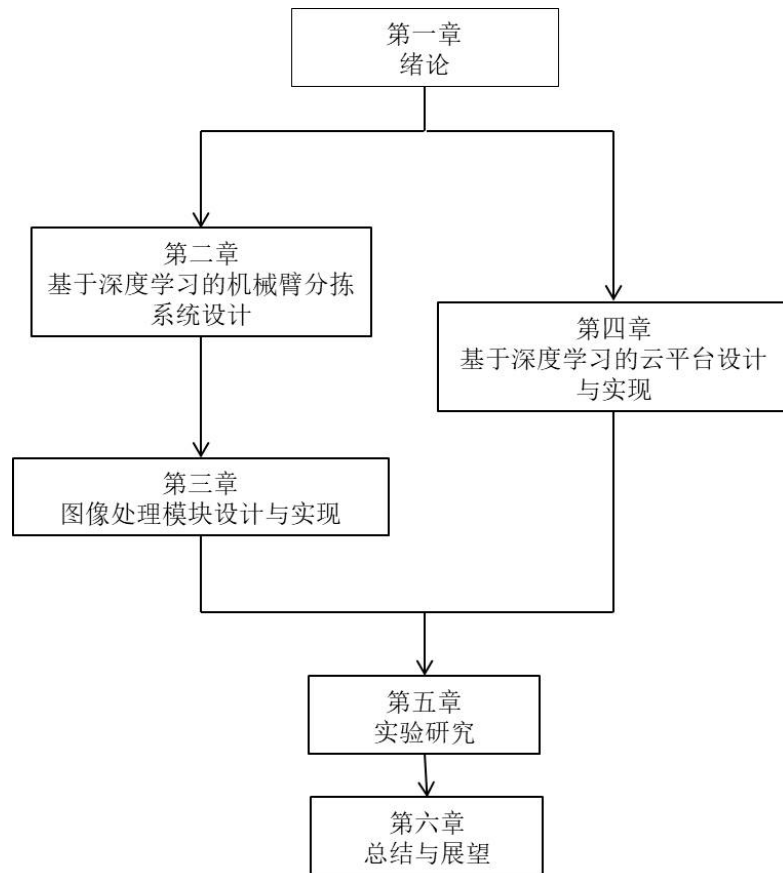


图 1-3 论文章节架构

1.4 本章小结

本章首先阐述了论文的研究背景和研究意义；然后介绍了目标检测算法、基于视觉的自动分拣系统和深度学习云平台的国内外研究现状；最后给出了本论文的主要研究内容和论文架构。

第2章 基于深度学习的机械臂分拣系统设计

自动分拣系统的核心为图像处理模块，该模块的核心为目标检测算法。传统的目标检测算法需要人为为工件构造特征，通过模板匹配的方式确定工件位置，再通过分类器的方式确定工件类别。当变换检测工件时，就需要针对工件再次构造特征。而基于深度学习的目标检测算法则可以实现端到端的训练。这使得基于深度学习的机械臂分拣系统具有很强的可移植性。

自动分拣系统是一个多模块协同作用的系统。包括图像采集模块、图像处理模块和机械臂控制模块。这三个模块均建立在硬件基础之上，且需要高效的通信机制。本章主要介绍整个自动分拣系统的整体架构和底层设计，包括硬件选型、通信流程等。然后介绍了如何选择本系统中使用的目标检测算法，至于具体模型的介绍、训练及部署等，将放到第三章去讨论。最后，本章将介绍目前相对已经比较成熟的相机标定及机械臂控制方法。

2.1 系统整体架构

系统的逻辑层面的架构图如图 2-1 所示。

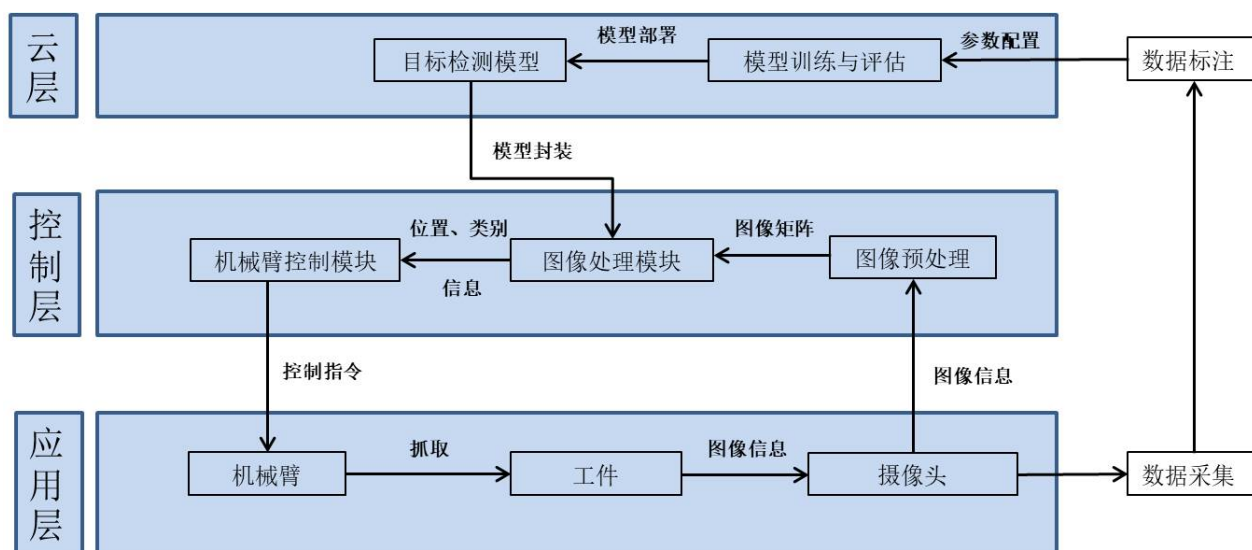


图 2-1 系统逻辑架构图

基于深度学习的机械臂分拣系统逻辑上可以分为三个层面：应用层、控制层和云层（即服务器层）。以下对每个层面进行具体介绍：

1. 应用层

应用层由机械臂、工件和摄像头组成。主要是摄像头收集工件信息，机械臂对工件执行抓取动作。这一层面是整个系统对外展示的层面，代表了整个系统的表现。

2. 控制层

控制层主要进行信息的处理和机械臂的控制指令生成。首先，控制层接收由应用层摄像头发送的图像信息，并进行预处理，之后发送给图像处理模块中的目标检测模型进行预测，得到工件的位置和类别信息，然后发送给机械臂控制模块，由机械臂控制模块生成相应的机械臂控制指令，发送给机械臂进行执行。

其中，摄像头和控制层的通信通过 USB 完成，所使用的摄像头为 USB 摄像头，主要是进行图像信息的传输；控制层和机械臂之间通过串口进行通信，主要是进行机械臂控制指令的传输，如 G 代码等；而控制层的模块之间，则通过机器人控制系统（Robot Operation System, ROS）进行，主要进行整型或浮点数的传输。

3. 云层（服务器层）

服务器层主要用于系统的搭建，在系统运行时是不参与工作的。服务器主要提供计算资源，用于深度学习模型的训练和评估。训练完成后，得到的其实是模型配置文件和由大量浮点数组成的权重文件。模型配置文件主要用于描述模型的结构和参数，权重文件则用于描述模型网络的各个参数值。有了这些，就可以将模型部署和封装到控制层，用于预测图像中工件的位置和信息。

此外，在三个层之外，还需要使用摄像头采集数据集，并进行标注，用作服务器层模型训练的数据集。

2.2 系统硬件选型

作为机械自动化及高性能计算一体的系统，基于深度学习的自动化分拣系统对硬件环境非常依赖。硬件性能的好坏将直接影响目标检测模型的迭代次数、机械臂响应速度和分拣的延迟。因此，自动分拣系统的硬件选型非常重要。

本节主要介绍系统的硬件选型，包括用于采集图像信息的 USB 摄像头、用于高性能计算的服务器、运行图像处理模块和机械臂控制模块的高性能嵌入式平台以及用于抓取工件的机械臂。

2.2.1 相机

2.2.2 目标检测模型训练用服务器

2.2.3 目标检测运行硬件平台

2.2.4 机械臂选型

2.2.5 系统硬件架构

2.3 系统模块间通信机制

2.3.1 USB 通信

本系统中，USB 通信主要用于摄像头向 TX2 传递图像信息。本课题使用的 USB 版本为 USB 3.0^[31] 版本，其传输速度为 5Gbit/s，足以支撑高 FPS 的图像传输。

2.3.2 ROS 通信

2.3.3 串口通信

2.4 目标检测算法选型

2.5 相机标定及机械臂控制模块设计与实现

第 3 章 图像处理模块设计与实现

3.1 图像算法理论

3.1.1 感知算法

3.1.2 YOLOv3

3.2 图像采集和数据获取

3.3 数据标注

3.4 目标检测模型训练与评估

3.5 目标检测模型部署

第 4 章 基于深度学习目标检测模型的云平台设计与实现

4.1 深度学习云平台方案设计

4.1.1 云平台整体架构

4.1.2 基于 Http 协议的通信方法分析

4.1.3 云平台性能需求分析

4.2 基于 Flask 的 Web 服务端设计与实现

4.2.1 云平台服务端架构

4.2.2 基于 Linux 的 Shell 脚本实现云平台模型训练

4.3 基于 Web 前端的客户端系统设计

4.3.1 客户端网页需求分析

4.3.2 客户端页面分析、设计与实现

4.3.3 Web 前端与服务器端的通信实现

第 5 章 实验研究

5.1 机械臂分拣系统性能实验

5.1.1 分拣延时实验

5.1.2 分拣准确率实验

5.2 基于深度学习的目标检测模型云平台性能测试

5.2.1 Web 前端性能测试实验

5.2.2 服务端性能测试实验

第 6 章 总结与展望

6.1 总结

6.2 展望

参考文献

- [1] 国务院. 新一代人工智能发展规划, 2017.
- [2] ISO. Standard 8373, 1994.
- [3] 陈见飞, 彭靖峻. 拉链齿数的数字图像处理研究 [J]. 南京: 工业控制计算机, 2011. 24(3):10–13.
- [4] 卞正岗. 机器视觉技术的发展 [J]. 中国仪器仪表, 2015. 06:40–42.
- [5] 张文辉. 基于机器视觉的机器人分拣系统研究 [D]. 西安: 西安建筑科技大学, 2018.
- [6] SJ Pan, Q Yang. A survey on transfer learning[J]. IEEE Transactions on knowledge and data engineering, 2010:1345–1359.
- [7] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Li Fei-Fei. Imagenet: A large-scale hierarchical image database[J]. Computer Vision and Pattern Recognition. CVPR, 2009:248–255.
- [8] Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. //F. Pereira, C. J. C. Burges, L. Bottou, K. Q. Weinberger, editors, Advances in Neural Information Processing Systems 25. Curran Associates, Inc., 2012. 1097–1105.
- [9] 方路平, 何杭江, 周国民. 目标检测算法综述 [J]. 计算机工程与应用, 2018:11–18+33.
- [10] David G. Lowe. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004. 60(2):91–110.
- [11] Yan Ke, R. Sukthankar. Pca-sift: a more distinctive representation for local image descriptors[C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. volume 2. 2004:I–II.
- [12] Bay H, Tuytelaars, GoolL. Surf:speeded up robust features[J]. Computer Vision & Image Understanding, 2006. 110(3):404–417.
- [13] Freund Y, Schapire R E. Experiments with a new boosting algorithm[C]//Proceedings of Thirteenth International Conference on International Conference on Machine Learning. 1996:148–156.
- [14] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 2014.
- [15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition[C]//David Fleet, Tomas Pajdla, Bernt Schiele, Tinne Tuytelaars. Computer

- Vision – ECCV 2014. Cham:Springer International Publishing. ISBN 978-3-319-10578-9, 2014:346–361.
- [16] Georgia Gkioxari, Ross Girshick, Jitendra Malik. Contextual action recognition with r*cnn[C]//The IEEE International Conference on Computer Vision (ICCV). December 2015.
- [17] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. //C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, R. Garnett, editors, Advances in Neural Information Processing Systems 28. Curran Associates, Inc., 2015. 91–99.
- [18] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. You only look once: Unified, real-time object detection[C]//The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 2016.
- [19] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg. Ssd: Single shot multibox detector[C]//Bastian Leibe, Jiri Matas, Nicu Sebe, Max Welling. Computer Vision – ECCV 2016. Cham:Springer International Publishing. ISBN 978-3-319-46448-0, 2016:21–37.
- [20] Joseph Redmon, Ali Farhadi. Yolo9000: Better, faster, stronger[J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017:6517–6525.
- [21] Joseph Redmon, Ali Farhadi. YOLOv3: An incremental improvement[J]. CoRR, 2018. abs/1804.02767.
- [22] 董腾. 基于机器视觉的水果分拣系统的研究 [D]. : 聊城大学, 2018.
- [23] 许江淳, 岳秋燕, 任向阳, 王晴. 基于机器视觉的药片表面缺陷识别与分拣系统设计 [J]. 传感器与微系统, 2017. 36(06):90–93.
- [24] 杨美程, 吴戈. 工件智能分拣平台设计研究 [J]. 电脑知识与技术, 2017. 13(08):178–179.
- [25] 伍锡如, 黄国明, 孙立宁. 基于深度学习的工业分拣机器人快速视觉识别与定位算法 [J]. 机器人.ROBOT, 2016. 38(6):711–719.
- [26] 陈翠. 你知道机器视觉在物流分拣自动化系统中是如何工作的吗? . <https://zhuanlan.zhihu.com/p/42018613>, 2018.
- [27] 何泽强. 基于机器视觉的工业机器人分拣系统设计 [D]. 哈尔滨: 哈尔滨工业大学, 2016.
- [28] alibaba. 阿里云. <https://www.aliyun.com/>, 2018.
- [29] Amazon. Amazon web services. <https://aws.amazon.com/cn/>, 2018.
- [30] NVIDIA. Nvidia gpu cloud. <https://ngc.nvidia.com/>, 2018.
- [31] Chih-Hung Huang, Chien-Cheng Kuo, Shih-Min Hsu. Universal serial bus (usb) 3.0 compatible host with lower operation power consumption and method for reducing operation power consumption of a usb 3.0 compatible host[Tech].

附录 A

这是附录 A 的内容。

附录 B

这是附录 B 的内容。

