

学校代码: 10246

学 号: 14210240056

復旦大學

硕 士 学 位 论 文  
(学术学位)

众包平台中图像标注任务的交互方法研究

Interactive Solution for Handling Image Annotation Tasks  
on Crowdsourcing Platforms

院 系: 计算机科学技术学院

专 业: 计算机应用技术

姓 名: 严鑫

指 导 教 师: 丁向华 副教授

完 成 日 期: 2017 年 3 月 30 日



## 指导小组成员名单

丁向华    副教授

张   亮    教   授

顾   宁    教   授

卢   噉    副教授



# 目录

目录.....	I
摘要.....	V
ABSTRACT.....	VII
第一章 绪论.....	1
1.1 研究背景、意义和问题.....	1
1.1.1 研究背景和意义.....	1
1.1.2 主要问题.....	2
1.2 国内外研究现状和相关工作.....	3
1.2.1 众包研究现状.....	3
1.2.2 图像标注任务研究现状.....	4
1.3 研究内容和方法.....	5
1.3.1 研究对象.....	5
1.3.2 主要研究内容.....	6
1.3.3 研究方法.....	7
1.4 本文的组织结构.....	8
第二章 图像标注任务的分类及平台搭建.....	11
2.1 图像标注任务调研.....	11
2.2 图像标注任务分类.....	15
2.2.1 语义标注.....	15
2.2.2 点位标注.....	16
2.2.3 区域标注.....	16
2.2.4 序列筛选标注.....	17
2.3 众包平台的搭建与配置.....	18
2.3.1 基础环境与框架配置.....	18

2.3.2 “众研”平台的搭建与介绍 .....	19
2.4 本章小结 .....	21
第三章 众包工作模式的对比性研究 .....	23
3.1 概述 .....	23
3.1.1 线下众包工作模式介绍 .....	23
3.1.2 线上众包工作模式介绍 .....	23
3.1.3 研究目的 .....	23
3.2 项目和任务介绍 .....	24
3.2.1 项目背景及介绍 .....	24
3.2.2 任务介绍 .....	24
3.3 对比性研究方法 .....	25
3.3.1 两种众包工作模式下的工作过程 .....	26
3.4 数据采集与分析 .....	27
3.4.1 数据采集 .....	27
3.4.2 定量和定性相结合的结果分析 .....	28
3.5 交互方法设计原理启示 .....	30
3.6 本章小结 .....	31
第四章 图像标注任务的交互方法设计与实现 .....	33
4.1 概述 .....	33
4.2 需求调研与问题分析 .....	33
4.2.1 任务上传过程介绍 .....	34
4.2.2 任务标注过程介绍 .....	35
4.3 设计原理与技术方案 .....	35
4.3.1 设计理论 .....	35
4.3.2 设计准则和方案 .....	35
4.3.3 交互模板方案设计概要 .....	36
4.3.4 关键技术和算法 .....	38

4.4 模块和界面展示.....	43
4.4.1 可视化功能优化.....	43
4.4.2 任务生成导入功能优化.....	44
4.4.3 标注交互界面展示.....	45
4.4.4 交互模板功能优化.....	48
4.5 本章小结.....	49
第五章 平台交互方法的评估 .....	51
5.1 实验设计.....	51
5.2 效率评估.....	53
5.3 用户体验评估.....	56
5.4 本章小结.....	58
第六章 总结与展望.....	59
6.1 论文总结.....	59
6.2 工作展望.....	60
参考文献.....	61
发表论文和科研情况说明 .....	65
攻读硕士期间发表的学术论文.....	65
攻读硕士期间参与的科研项目.....	65
致 谢.....	67





## 摘要

随着互联网技术的发展,完成一些常见或者新型任务的工作方式变得更加多元化。越来越多的人工智能任务每天都在线下或者众包平台上被完成,其中一类就是图像标注任务。通过这种渠道可以为机器学习领域快速的提供大量的已标注数据,作为它们模型和算法的训练集,可以有效的帮助机器学习中模型和算法的优化。

目前主要有两种众包工作模式支持完成图像标注任务,一种是较为传统线下众包工作模式(通过即时通讯软件+相应的图像辅助软件+文本记录软件等),另一种是通过线上众包平台来完成,例如:亚马逊土耳其机器人, LabelMe, 数据堂和中国众宝网。众包-通过将工作外包给网络大众的方式,提供了一种更加经济有效的生产方式。目前线上众包平台基本支持最简单的图像识别标注任务,例如图片中是否有猫、有车或有人,在交互方法上也基本上已经标准化。然而,对稍微复杂的图像标注任务,例如标注图像中身体的各个部位,当前的众包平台在支持这些任务的快速发布以及标注的可用性和有效性方面仍然缺乏一套有效的交互方案。那么究竟有多少种图像标注任务需要被解决?怎样去更好的支持它们快速的上传到众包平台?如何改善用户(尤其是残疾人等弱势群体)的标注体验并提高工作效率?针对解决这些问题,本文的研究工作主要包括以下几方面:

- (1) 通过对当前众包平台上进行的图像标注任务进行统计,然后从机器学习领域和图像识别领域中对已标注图像数据的需求进行相关调研,结合二者,对当前主要需要被解决的图像标注任务进行抽象和分类。在本文中,将其分为了四类,分别是:图像语义识别任务,图像点位标注任务,图像区域标注任务,图像序列化筛选任务。
- (2) 首先搭建众包原型系统“众研”,展开了对比性研究,即通过两种不同的众包工作模式来完成实际的图像标注任务,一种是传统的线下众包工作方式(通过即时通讯软件+相应的图像辅助软件+文本记录软件等),另一种则是目前流行的众包平台的工作方式,通过对比分析得出了利用众包平台在解决图像标注任务时的优缺点,为本文后续对交互方法的设计和原型系统的优化提供了设计原理启示。
- (3) 基于用户参与设计式理论,并结合之前对比性研究得出的相关设计准则,对分类得出的四类图像标注任务进行了交互方法和算法设计,生成一对一的交互模板,并对相关功能模块进行了优化,最后,通过实验数据验证本文设计

和实现的交互方法能有效的提升图像标注任务的标注效率和用户体验。

**关键词：**众包；图像标注任务；在线众包平台；人工智能任务；众包工作

**中图分类号：**TP3

## Abstract

With the development of internet technology, it's more diversified for the way to accomplish the ordinary or new tasks. A growing number of human intelligence tasks (HITs) are completed every day on online crowdsourcing platform or offline, and one type of such task is image annotation. This channel can provide a large number of annotated data as the training sets for the machine learning area to optimize their models and algorithms.

There are many crowdsourcing platforms such as Amazon Mechanical Turk, LabelME, DataTang, and ChinaCrowds that support image annotation. By outsourcing work to a network of people, crowdsourcing provides a more economical and efficient means of production. However, there is still a lack of integrated solutions for image annotation. How many kinds of image annotation tasks are there to be solved? And how can they be better supported? In order to solve these problems, the main work and contributions are threefold:

- (1) We first did a survey on image annotation tasks in the machine learning and image recognition fields and the current crowdsourcing platforms. And then, we made a classification of these image annotation tasks, we have classified four types of image annotation tasks: semantic annotation, point annotation, region annotation, and sequence annotation.
- (2) We built a prototype system called ZhongYan as our experimental crowdsourcing platform and conducted a comparative study of crowd work with and without crowdsourcing platforms, trying to understand the advantages and disadvantages of using crowdsourcing platforms. And then, we summarize some design principles based on our study and the design concept of social translucency. This principles is benefit for the design of image annotation tasks' interactive method and the optimization of the prototype system.
- (3) Based on the concept of user participation in design and the relevant design criteria that obtained from the previous comparative study, design and implementation of the interactive interface for the four types of image annotation tasks. And finally, through the experimental data to verify the interactive method can improve the efficiency of image annotation and user experiences.

**Key words:** Crowdsourcing; Image annotation tasks; Online crowdsourcing platform; Human intelligence tasks; Crowd work

**Chinese Library Classification:** TP3

# 第一章 绪论

众包任务种类多样并且可以通过多种工作方式完成,其中一类就是图像标注任务,目前众包平台在解决图像标注尤其复杂的图像标注任务的交互方式中存在着诸多不足,从而限制了任务发布和完成的效率以及用户体验;因此,迫切需要设计和实现一种基于众包平台的图像标注任务的交互方法,用来改善将图像标注任务上传到平台以及在平台上完成图像标注这两个过程中的效率和用户体验。

## 1.1 研究背景、意义和问题

### 1.1.1 研究背景和意义

2015 年李克强总理提出的“互联网+”的概念已经融入到各行各业[45],促使许多行业数据量的暴增,在这样一个大数据时代下,通过合理的方法去分析和使用这些数据可以有效的帮助我们解决许多的问题。机器学习领域通过收集有效的标注数据去训练和优化它们的模型和算法,从而可以更好的对一些事物和问题做出决策和推论;在整个过程中,研究人员首先通过某些算法指导计算机利用一些已知的数据去训练出适当的模型,然后利用得出的模型去对新的情境给出判断,为了获得一个较好的模型,不仅需要足够好的算法,还需要大量的已标注数据(例如对图像、视频、文本等数据类别/内容的标注数据)作为训练和分析的基础,因为在样本训练的过程中,拥有足够多的已标注数据集,可以不断有效的改进相应的算法和模型,从而使其更加可靠;但如果仅依靠科研人员来对这些庞大的数据进行标注,去生成他们所需的训练数据集,那么工作量显得非常之大,另外,制造已标注的数据集不需要专业的技术背景,具有基本认知能力的普通人便可胜任此项任务。

随着生活水平的提高以及互联网时代的发展,网络覆盖到了越来越多的地方,越来越多的网民通过各种类型的电脑或移动终端设备浏览网络。正是在这样的背景下,近些年来涌现出一种新的生产方式-互联网众包,即让网民一起参与到问题的解决中来。因为网民群体的数量较多且响应较快,所以互联网众包模式以其快捷、低成本的特点,为解决大数据标注工作提供了一种友好的渠道。同时,这也能为许多弱势群体(例如残疾人,老年人)提供较多的就业机会,使得他们能够通过网络争取一定的酬劳。

2006 年, Howe J.[1]首先提出了“Crowdsourcing”(众包)这一概念,并认为众包是一种新兴的工作模式,是指公司、机构或者个人把过去由特定员工执行的工作或者任务,指派给非特定的大众群体的一种形式。近年来,随着互联网的发展,越来越多的企业和科研机构开始尝试将具有一定创新性和技术性的工作任务通过互联网渠道委托给外部个体或组织完成,基于互联网并且注重开放式协作,这些都是众包模式的特点[2]。一般而言,众包模式包括三个主体,分别是:有任务需求的发包方,有意愿完成任务的接包方和连接二者的平台,这其中平台主要指代众包平台,但是也包括一些使得二者能够顺利完成任务交接的其他媒介,例如在线社区、论坛、即时通讯软件等[3]。

所以,众包平台以一座桥梁的形式连接着任务需求方和工人。Amazon's Mechanical Turk(简称 AMT)就是一个比较有名的众包平台[27]。在这个平台上:任务需求方通常是一些研究机构或科研人员,工人通常是一些在网络上通过完成任务获取报酬的一些众包工作者;平台可以帮助这两类用户去管理并完成一些暂不适合计算机执行的任务,这类任务在 AMT 上被定义为人类智能任务(Human Intelligence Task),例如,图像/文本/声音/视频等识别和标注任务等,相比机器,因为人本身有足够的背景知识和良好的归纳能力的优势,所以在面对这些任务的时候(特别是在机器学习过程中,需要大量高质量的训练数据时),众包模式凭借其网络大众群体的大规模优势,能够以足够快并且成本较低的方式来完成这些任务。

### 1.1.2 主要问题

通过对当前众包平台以及图像标注任务的研究,发现在众包平台上完成复杂图像标注任务的工作过程中存在以下几个明显的主要问题:

- (1) **任务发布周期长:** 任务需求方将图像标注任务发布到众包平台上的耗时较长。
- (2) **缺乏一个系统的分类:** 当前的众包平台中适配的图像标注任务的种类少且较为基础。无法满足机器学习领域对于各种图像识别研究的已标注数据的采集。
- (3) **标注效率低:** 完成图像标注任务的过程中,容易出错,且标注速率较慢,导致完成任务效率低。
- (4) **用户体验差:** 图像标注任务标注过程和上传过程比较复杂,界面标注过程不够简练,导致平台的两类用户(标注者和任务发布者)缺乏一个良好的用户体验。

综上所述, 本文通过对众包平台中图像标注任务的交互方法进行研究, 做出了以下设计和改进, 从而使得在众包平台上能够更好的完成图像标注任务, 主要工作如下:

- (1) 通过对当前图像标注任务的解决方案以及当前机器学习领域所需要的图像标注任务的调研分析, 抽象出这些任务的特性, 并对其进行了详细的任务分类。方便图像标注任务的支持任务发布方的个性化需求, 能够较为快速地发布图像标注任务到众包平台上。
- (2) 对已有的图像标注任务类别分别构造相应的模板, 并且允许用户设计和上传自己任务演示模板, 设计了一个可扩展且能够动态更新的任务模板库。
- (3) 以用户体验为中心。在任务交互界面的设计上, 要以用户为出发点, 降低用户对系统的认知差距, 提高可操作性, 从而提升任务完成效率和准确率。

最后通过实际图像标注项目的测验以及平台交互方法的评估, 证明本文基于众包平台设计的图像标注任务的交互方法能够有效地解决当前的众包平台在完成图像标注任务过程中所面临的几个问题, 主要体现在以下几点: (1)促使任务需求方能够方便、快速且批量式的发布任务到众包平台上; (2)通过改善标注界面设计, 提高了众包工作者在众包平台上完成图像标注任务的工作效率以及用户体验。因此, 本文针对众包平台中图像标注任务而设计的交互方法具有重要的研究意义。

## 1.2 国内外研究现状和相关工作

### 1.2.1 众包研究现状

当前围绕众包工作, 有许多比较接近的术语, 例如: crowdsourcing, collective intelligence, human computation, serious games, peer production[4,5,6,7]。本文关注的是众包工作请求者(个人, 团体, 组织)给予支付的在线众包工作。概况来讲, 众包工作是围绕一组连接着组织、个体、技术和工作的关系而形成的社会技术工作系统[8]。

围绕众包工作, 科研人员也展开了比较全面的研究, 当前许多研究的注意力都集中在如何在技术上实现自动化、改善工作过程、质量控制等, 例如, 提供测试来保证质量以及任务自动分配等[3]。通过在工作人员中选出一个最佳工作人

员来管理、教导和监督其他工人来实现更好的质量管理[9]。面向残疾人的众包研究主要体现在无障碍性[10]。例如加强盲人等公共交通乘客的独立性和安全性[11]，以及提高残疾人对众包平台的访问性[12]。

众包在国外(尤其是欧美等发达国家)已经发展的比较成熟，很多的公司和企业通过众包模式去经营和管理自己的业务[23]，这也使得众包理念渗透到了国外的各行各业中，利用众包的开放创新以及群体智慧的特点，帮助他们保持一个良性的发展趋势。在企业发展过程中，也诞生了许多优秀的众包平台，例如：科学问题平台 InnoCentive[25]、摄影照片分享平台 iStockphoto[26]、词条百科平台 Wikipedia[27]、帮助解决人类智能任务的 AMT 平台[28]、全球网上工作平台 upWork[29]以及数据处理平台 CrowdFlower[30]等。

相比较国外较为成熟的众包模式，国内还处于一个初级阶段[14]，在一些需求量大的行业(如设计业，网站开发业等)发展较为迅猛，但是其他行业与国外的一些众包企业仍存在一定的差距。国内也有许多的众包平台，规模较大且比较有名气的主要有 K68[13]、猪八戒[15]、时间财富网[16]、互帮网[17]、任务中国网[18]和耳目网[19]，其中互帮网和耳目网主要面向的众包工作者类型为残疾人，旨在帮助残疾人增加收入机会；另一方面，目前国内面向大数据标注的众包平台：如针对大数据共享交易的众包平台数据堂[20]，众宝网[21]等，这些平台承担了国内许多的众包工作。

但是经过我们对一些残疾人众包工作者的研究后发现[22]，有许多的众包任务仍然是采用一些较为传统的众包工作模式完成的。例如我们采访的一些残疾人团体：任务方通过线上 QQ 与他们进行交流，先达成协议，然后线下再通过任务需求方提供的图像标注软件来完成任务需求方传输给他们的图像标注任务。

## 1.2.2 图像标注任务研究现状

当前市场上有一些图像标注软件，传统的众包工作模式也主要依靠这种软件去获取图像数据的标注结果，常见的有：Evernote Sketch[31]，Image Annotation[32]等。这类图像标注软件的优点在于：功能较为齐全，可以方便的对本地图片进行标注。但是缺点也很明显：由于软件开发周期长，导致无法快速适应标注数据的格式需求变化，另外由于移植性差，无法直接投入到众包平台上使用，回收标注结果也比较麻烦。

概括来讲，这种传统的软件标注方式的缺点主要体现在以下三个方面，(1)大数据多次传输消耗较多时间；因为需求方需要将待标注的数据传输给这个工作



团体的负责人,再由负责人进行任务分配然后再依次传输任务数据给团体中的每个人,导致数据多方周转,消耗大量时间。(2)应对新的标注需求,响应周期过长;因为当新的标注需求到来时,首先需求方要进一步开发标注软件,重新发布,然后每位工作者重新进行下载,安装,使用新的标注软件,才能继续完成标注任务。(3)完成任务的过程不方便,导致时间相对使用众包平台较长;因为使用众包平台可以简化一些标注过程的操作,从而降低完成任务的时间。

现有的众包平台中,MIT 提供的 LableMe 平台[33]使得用户可以对图像进行标注,它的目标是为计算机视觉领域提供一个在线的图像标注工具,用来构建已标注图像数据库,方便科研人员的数据获取。它主要解决的是对象的轮廓标注,通过在图像中绘制对象轮廓的方式,来对物体进行标注,优点在于能够精准的确定对象的位置,但对于普通的众包工作者而言,并不是很容易快速上手,而且任务标注时间较长,导致效率较低,无法快速获取劳动报酬。另外,Amazon's Mechanical Turk 平台(简称 AMT)上发布的人类智能任务,有许多是关于图像标注任务的,但是只适配了一些较为简单基础的图像标注任务和交互方法,例如:给出一张图片,通过问答的形式,标注它所属类别或其他一些信息,显然这不足以解决一些复杂的图像标注任务(例如:要求对人物图像的各个身体部位进行精确的像素坐标位置等)。国内众包平台中的众宝网以及数据堂 DataTang 上也支持一些图像标注任务,但是大体的形式和 AMT 类似,只针对一些简单基础的图像语义标注任务。

## 1.3 研究内容和方法

### 1.3.1 研究对象

本文以当前众包工作模式、图像标注任务、众包平台、任务需求方、众包工作者以及图像标注交互方法作为主要研究对象,

- (1) **众包工作模式**,是一种新兴的工作模式,是指公司、机构或者个人把过去由特定员工执行的工作或者任务,指派给非特定的大众群体的一种形式。具体来讲,任务方通过在线社区、论坛、众包平台或即时通讯软件等媒介,邀请或雇佣众包工作者来完成相应任务。
- (2) **图像标注任务**,一般指用户在原有图像信息的基础上需要获取一些其他与图像本身相关的其他数据,从而满足自己的研发需求。所以会通过任务的形式指派给特定的人员,按照自己的要求帮助完成标注,从而获取额外的图像信息。

- (3) **众包平台**，是一种新型的商业化网络平台，也即公司、机构或个人利用网络众包平台将自己的技术方案设计或其他的工作任务发放出去，利用众包平台下庞大的用户群体来辅助自己完成这些任务。
- (4) **任务需求方**，一般指代拥有一些工作任务的公司、机构或者个人。
- (5) **众包工作者**，指代一些社会个体，他们利用自己的业余时间，依靠众包工作模式完成一些任务需求方的制定任务，从而获取报酬。
- (6) **图像标注交互方法**，一般指代辅助用户顺利完成图像标注任务的一套交互方法，包括图像任务的上传、标注等过程。

### 1.3.2 主要研究内容

针对上述所提到的现有的完成众包任务中图像标注任务的两种工作方式的不足，结合后续的一些需求分析，基于所定义的设计原则下，设计一套基于众包平台的图像标注任务的交互方法来针对性的解决上述所提的主要问题。如图 1.1 所示：

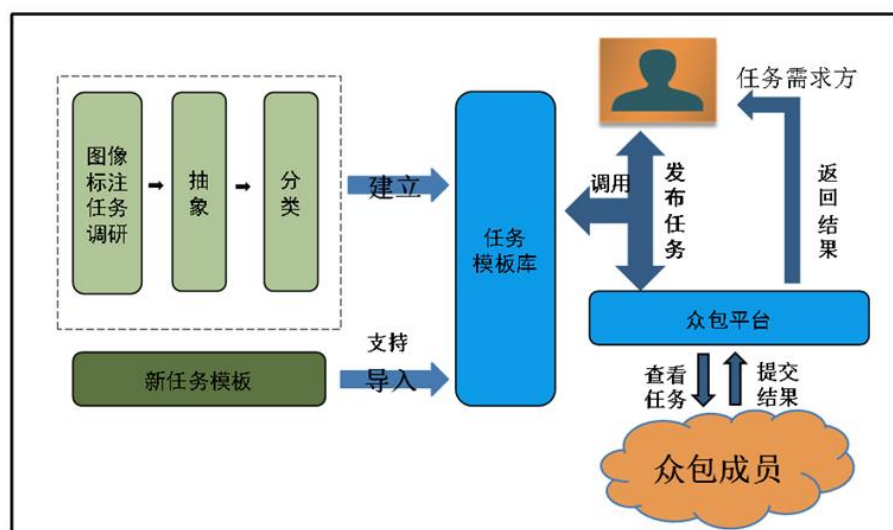


图 1.1 研究内容

任务需求方能够使用众包平台导入任务数据。一般情况下，任务方将任务托管给平台工作人员，由工作人员代理导入任务数据，为方便任务方自主的上传任务数据，平台应支持这一功能。应对常见的图像标注任务，建立通用模板，能够动态更新，需求方可自由定义任务输入和输出格式。优化任务标注界面设计，改善标注任务的交互过程，使得用户能够快速，准确的完成标注任务。

### 技术路线

- (1) 构建图像标注任务分类。

- (2) 搭建众包实验平台。
- (3) 展开两种众包工作模式对比性研究，分析设计准则。
- (4) 设计和优化任务上传方式，方便标注数据的快速导入。
- (5) 建立图像标注任务模板，进行任务界面中输入和输出格式的交互设计。
- (6) 优化任务标注交互方法。

### 1.3.3 研究方法

本文将采用定量和定性相结合的分析方法研究这些要素间的内在关联，如图 1.2 所示。

#### 数据采集方法

为了对当前众包工作模式中图像标注任务的交互方法进行比较全面的深入分析，本文将使用多种方式对数据进行采集，通过定性和定量相结合的方式，从而更好的面向众包平台中图像标注任务的交互方法进行研究并提供设计依据。

利用众包工作模式来解决图像标注任务的过程中，用户行为数据主要有以下几类：第一类是标注数据，指代和标注任务相关的数据；第二类是日志数据，在众包平台系统中产生的系统日志；第三类是实证数据，包括通过用户访谈，调查问卷，参与观察等方式获取的一系列实证数据。

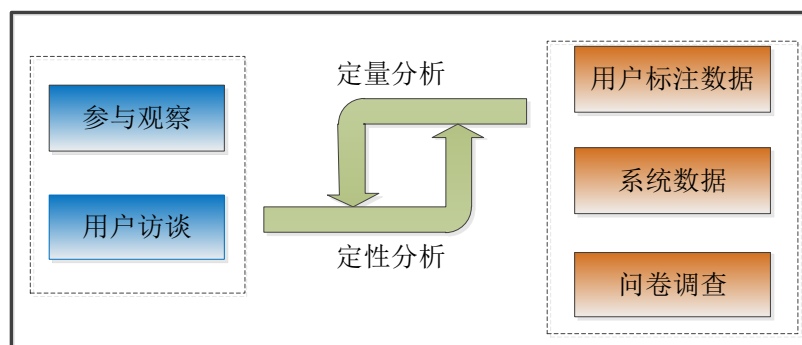


图 1.2 定性和定量相结合的分析方法

#### 用户标注数据

标注数据是指用户在完成图像标注过程中产生的一系列数据，是分析用户标注行为的主要数据之一。在不同的众包工作模式下，用户产生的标注数据形式不同。在线下众包工作模式下，需要用户在标注过程中配合记录相关数据，从而完成采集；在线上众包工作模式下，则一般通过众包平台系统记录相关数据到数据库，从数据库中完成采集。

## 系统数据

系统数据一般指代用户在众包平台中所产生的一系列有记录的数据。本文采集的系统数据包括两种，一种是用户在众包平台中产生的日志数据，主要包括用户使用众包平台的 IP，时间，功能，请求页面，请求动作等相关的统计信息。另外一种是在用户在众包平台中产生的内容数据，主要包括用户创建的项目信息，任务信息，标注结果信息以及发表的内容等。

## 参与观察

参与观察法(Participant Observation)是一种定性研究方法，指的是研究人员要融入到调查对象的日常生活中去，通过参与和观察他们的社会生活过程，随时随地的将一些相关数据记录下来。这种方法现在被广泛用于各个学科领域[36]，主要包括人种学，社会心理学等。本文中通过参与观察的形式了解任务需求方和众包工作者在完成图像标注任务过程中所呈现出来的特征以及认识和体会。

## 问卷调查

问卷调查是研究人员针对一些特定问题，在对象群体中展开的意见调查方法，是社会调查中的一种有效的数据采集手段[37]。结构化、合理化的设计和编写问卷调查中的一系列问题和可供选择的答案可以有效的帮助研究人员进行更快速的数据统计与分析[38]。在本文中，主要使用调查问卷来调查用户的一些基本信息和相关的标注数据，以及用户在标注过程中的体验及认识等。

## 访谈

访谈也称自由交谈，是人种学研究方法中的一种，常用来作为定性分析研究。一般由两人或多人对话构成，访谈者通过与对话者的问题交流过程去获取想要的一些信息[39]。因为访谈可以面对面的观察到被采访者的面部表情，肢体动作等，结合这些信息，可以更好的理解被采访者所表达的意思，这是访谈的优势所在。这种方法在许多领域都被广泛使用，主要包括人机交互，计算机协同工作领域等[40]。在本文中，通过对用户的访谈去了解一些比较主观性、不易简单统计的问题信息，从而丰富和补充其他数据。

## 1.4 本文的组织结构

本章针对机器学习中数据标注任务在计算机信息技术发展背景下的新需求，归纳了当前在互联网众包模式下现有的工作方式在完成图像标注任务中的不足，

并基于这些问题对国内外现有的图像标注软件以及众包平台进行了调研,结合相关文献提出了一套基于众包平台的图像标注任务的交互方法为上述问题提供了一种切实可行的实践解决方案。

本文主要分为六个章节(各个章节的组织结构如图 1.3 所示):

第一章,绪论:主要介绍本文的研究背景和研究意义,以及在众包模式中解决图像标注任务相关的研究现状和工作。同时还介绍了本文的研究对象、研究内容、预期技术路线以及研究方法。

第二章,图像标注任务的分类及平台搭建:主要介绍交互方法设计与实现的相关准备工作。包括:对于图像标注任务的调研分析、抽象和分类,以及自己搭建的众包平台-“众研”的相关过程介绍。

第三章,众包工作模式的对比性研究:主要通过搭建的众包实验平台“众研”针对这两种众包工作模式进行了对比分析并得出一些结论,将此作为我们设计图像标注任务交互方法的原理和准则。

第四章,图像标注任务的交互方法设计与实现:主要介绍基于前面总结的设计准则和需求分析,针对“众研”平台上图像标注任务标注过程进行的交互方法的设计与实现。

第五章,平台交互方法的评估:主要介绍从多个方面针对“众研”平台中图像标注任务而设计的一套交互方法进行的相关评估工作。

第六章,总结与展望:在全文的基础上,对主要完成的研究工作进行总结,并根据现有研究存在的不足和问题,然后对未来的工作进一步提出展望。

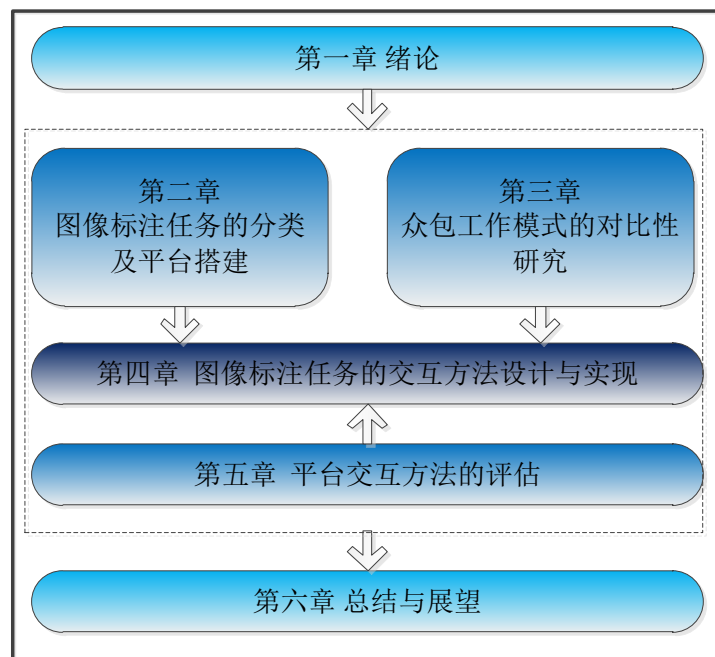


图 1.3 论文各章节的组织结构

## 第二章 图像标注任务的分类及平台搭建

因为实际的图像标注任务比目前众包平台上的一些基础性的标注任务要复杂很多,所以需要对图像标注任务进行一个比较通用的分类,然后按照类别针对性的提出相应方案;首先需要在机器学习以及图像识别等科研领域和当前众包平台上对于图像标注需求和具体的运作情况,进行一个详细的调研,通过对这些需求的详细分析从而将这些图像标注任务的特性抽象出来,从而进行具体的分类;另外还需要构建了一个众包平台,作为我们针对图像标注任务研究的开发、设计和评估的实验平台,本章主要介绍这些准备工作。

### 2.1 图像标注任务调研

图像标注在很多应用领域均有涉及,例如:日常应用领域,图像识别和机器学习等科研领域;总体来讲是为了更方便的进行图像检索以及更好的为研究人员提供有效的已标注数据集。

在日常生活中涉及到的主要是一些简单的图像标注任务,并且大多数是围绕“分类”这一共性来对图像进行标注,例如:在处理个人照片时,为了能够对一堆图片进行归类整理,人们会直接对图片打上标签,将相同标签的图片归在一类;又如:在一些图片社区共享网站上(如 Flickr, TrekEarth 等),许多用户会上传一些自己的图片,上传后会添加一些标注词用来概括这些图片的语义,比如拍摄图片的地点,图中的人物和景物等,因为这些社区拥有足够多的用户群体和已经标注好的大量图片,通过相似图片的比对能直接为后续上传到社区的图片推荐有效的标注词,这是图像标注的一种有效应用。

在图像识别和机器学习领域中为了更加智能的解决一些复杂的问题,例如:以图搜图,确认图像中某个物体位置,无人车自动驾驶,让机器读懂图片和视频等具体问题上,则需要大量的已标注图片数据作为训练集,通过不断的改进学习算法和模型,才能更好的解决这些问题。根据问题的需要,它们对于图像的具体标注信息需求也不同:有的研究需要图像中某些物体的位置或区域信息、有的需要对视频中的一些特定异常的图像帧进行挑选。图像识别的发展过程是一个从简单到复杂的识别过程,现在得益于计算机处理速度的提升以及大量已标注数据作为数据训练集的支撑,通过机器学习可以处理更复杂的图片识别任务。

图像标注,可以分为人工标注和自动标注两类。图像自动标注一直是国内外

研究重点,但是受限于自动标注的精度和准确性,还不能直接为机器学习和图像识别等领域提供大量有效的已标注数据。使用人工方式进行图像标注是比较直接和有效的方式,这种方式为科研领域提供了很多的训练集。例如:AMT 平台上支持许多科研工作者发布人类智能任务,包括一些简单基础的图像标注任务,通过支付酬劳的方式,吸引了许多的众包工作者参与这些任务,去完成标注;另外麻省理工学院(MIT)提供了一个在线的标注工具(LabelMe),用户在网络上使用这个工具,可以对图像进行标注,它的目标是为计算机视觉领域提供一个在线的图像标注工具,用来构建已标注图像数据库,方便科研人员的数据获取。它主要关注对象的轮廓标注,通过在图像中绘制对象轮廓的方式,来对物体进行标注,优点在于能够精准的确定对象的位置,但对于普通的众包工作者而言,并不是很容易快速上手,而且任务标注时间较长,导致效率较低,无法快速获取劳动报酬。

本文对当前国内外众包平台(Amazon Mechanical Turk, LabelMe, 数据堂, 众宝网等)的图像标注任务做了相关的调研,常见的图像标注任务种类如下:

## (1) 图像标记

如:给出一张图片,给图像定义一些标签,如图 2.1 所示

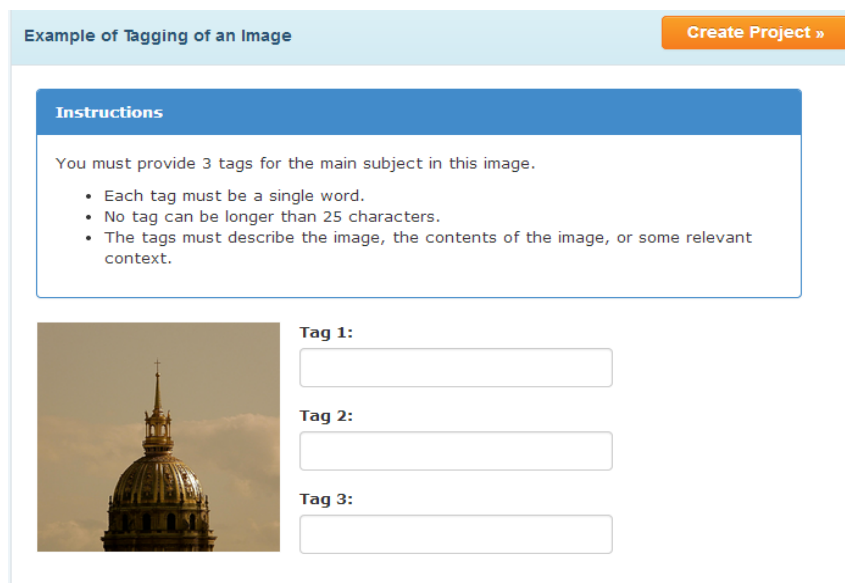


图 2.1 图像标记任务

## (2) 图像语义标注

这种类型的图像标注是在上述类型(图像标记)上的扩展,一般形式是提供一张图片,然后提出问题,比如一些图像转录任务,如图 2.2 所示,甚至提供一些可选答案,用户可以通过单选框或者文本框来提交答案,从而进行标注。如图 2.3, 2.4, 2.5 所示:



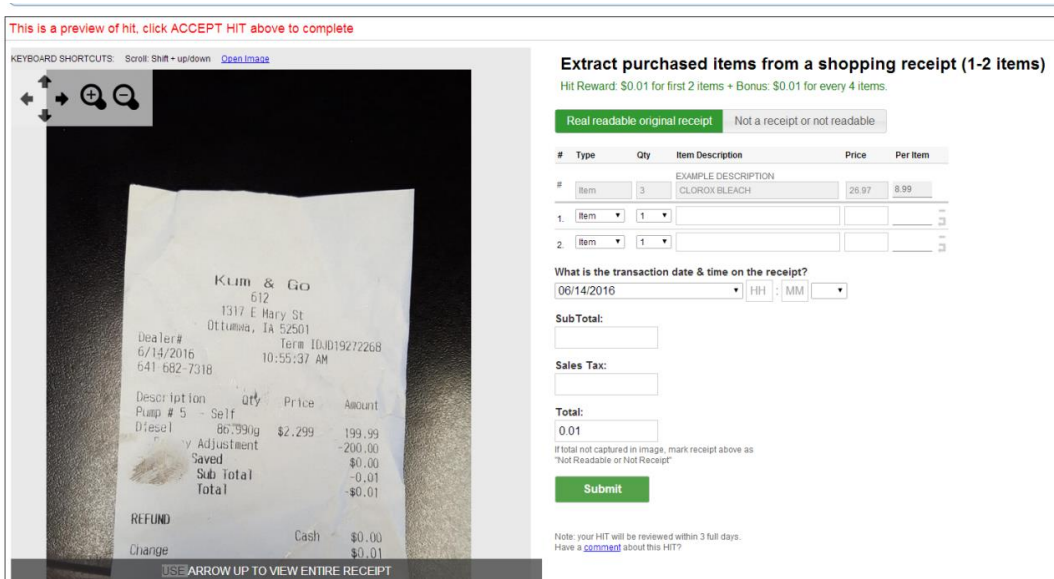


图 2.2 图像转录任务

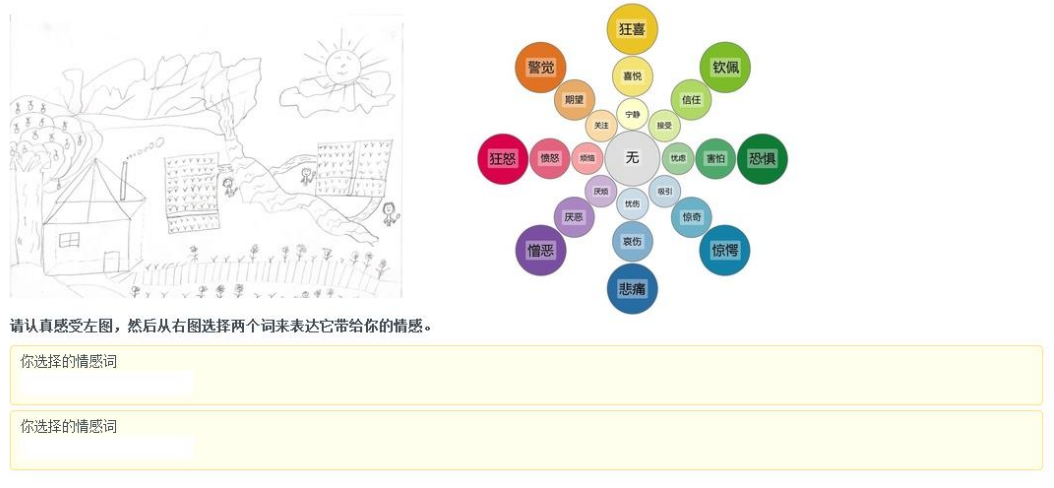


图 2.3 图像情感标注任务

(Q4) Does the blue box on the page contain a complete list, without any additional content?

☐ Yes ☐ No ☐ I am not sure

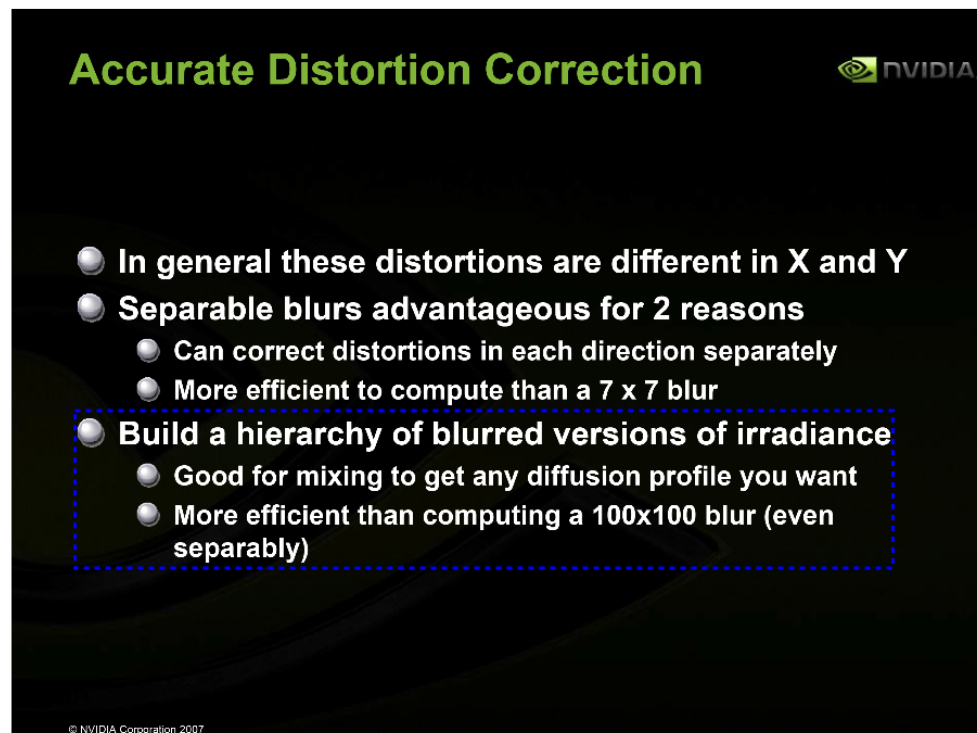


图 2.4 图像内容问答

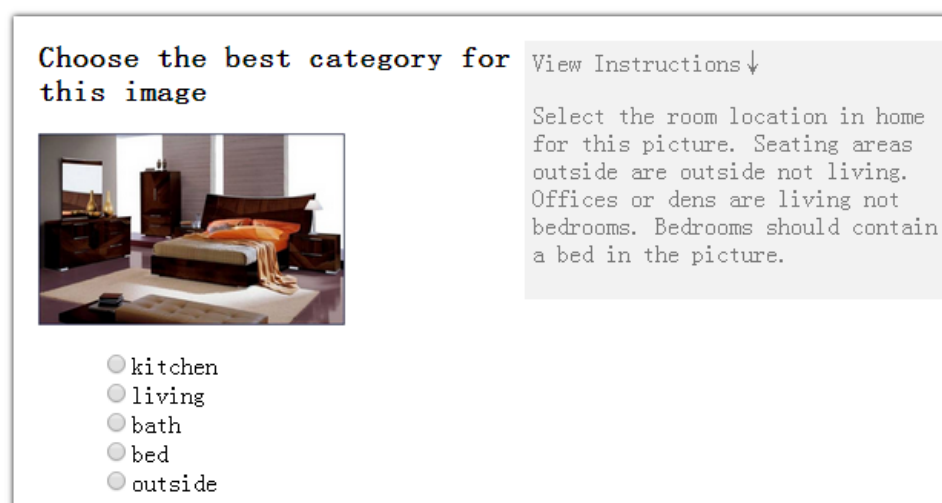


图 2.5 图像分类

### (3) 图像物体区域标注



图 2.6 物体区域标注

总得来讲，当前众包平台给出的图像标注任务形式较为简单，缺乏一个完整详细的图像标注任务分类。

## 2.2 图像标注任务分类

综合上述调研与分析，本文将当前涉及到的图像标注任务分为以下四类，分别是：语义标注，点位标注，区域标注，序列化筛选标注。其中，序列化筛选类别的任务主要面向多张连续图像的标注，其他三类则主要面向单张图片的标注。

### 2.2.1 语义标注

图像语义标注，主要针对单张图片的语义概括或者标签定义，例如：为了比较直接，快捷的对图片的整体信息有一个了解或对其进行归类，需要对其添加一至多个关键词进行标注。或者通过绑定一个问题的形式提供给众包工作者进行标注，目前很多的标注平台适配了这一类别的图片标注形式。如图 2.7 所示(描述图片)：



图 2.7 描述图片

### 2.2.2 点位标注

图像点位标注，主要针对单张图片中的多个部位确定，对该图片中需要确定的部位的位置进行标注，以方便用户能够快速定位到图片这些部位的位置。该类图片标注需要提供待标注部位的名称集，众包工作者根据名称集在该图片中标注出具体的位置点。如图 2.8 所示(确定图片中多个部位的位置点)：



图 2.8 确定图片中脸部多个穴位的位置点

### 2.2.3 区域标注

图像区域标注，主要针对单张图片中的物体识别，并确定物体所在的具体区域，需要对该图片中所需识别物体的位置区域进行标注，以方便用户能够快速定位到图片这些部位的位置，并且能够从图片中截取这些物体，进行后续的分析。该类图片标注需要提供待标注物体的名称集，众包工作者根据名称集在该图片中

标注出具体的物体区域。如图 2.9 所示(确定图片中的交通标识的位置区域):



图 2.9 识别图片中的交通标识

## 2.2.4 序列筛选标注

序列筛选标注,主要针对多张图片或视频的图像帧序列,通过对多张图片进行筛选,标注记录下某一些特殊图片的序号,以方便用户能够快速定位到图片集合中这些特殊图片的序号位置。该类图片标注需要提供详细的筛选标注规则,众包工作者根据筛选标注规则在该图片中标注出具体的图片序号。如图 2.10 所示(确定唇语图片中数字序列的起始位置)。

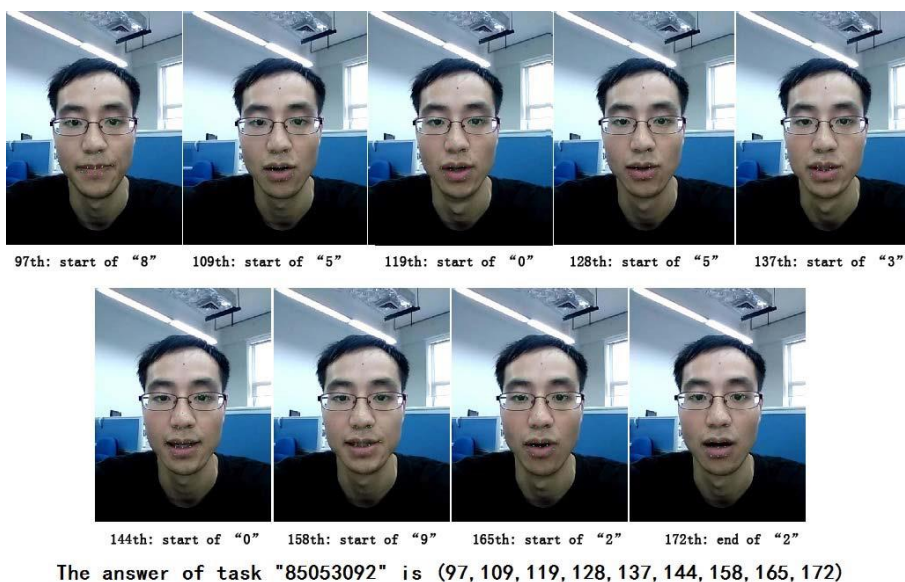


图 2.10 确定数字序列的起始位置



## 2.3 众包平台的搭建与配置

在现有的众包平台框架中, pybossa[34]是一套开源的众包框架, 用于分析或者丰富不能由机器单独处理的数据, 支持一些基本的智能任务的发布和管理, 在此框架基础之上, 本文通过二次开发, 搭建了自己的原型系统, 取名为“众研”, 以便于后续对图像标注任务的交互方法进行研究。本小结将介绍基础环境与框架配置以及“众研”平台。

### 2.3.1 基础环境与框架配置

#### a. 安装 Linux 系统

本文实验使用的是 Ubuntu 14.04.2 LTS, Ubuntu 系统可以在 VMware Workstation 或者其他虚拟机软件中安装。虚拟机软件的特点在于: 它可以模拟众多操作系统(如 Windows、Linux 和 Mac OSX 等), 操作过程与真机类似。主要安装过程如下:

- (1) 下载并安装 VMware Workstation。
- (2) 下载 Ubuntu14.04.2 LTS 镜像文件。
- (3) 在 VMware Workstation 下按照以下步骤操作: 新建虚拟机, 然后安装来源选择: Ubuntu14.04 镜像文件, 输入简易安装信息(用户名, 密码), 设置内存和虚拟硬盘大小, 创建好虚拟机后点击完成。
- (4) 安装完成后, 则会显示用户登录界面, 输入之前设置的用户名和密码便可进入 Ubuntu 的图形界面。

#### b. 相关软件和库文件的安装

搭建 Pybossa 框架需要的软件和库包括: git, PostgreSQL, python-virtualenv, Redis 等。

python-virtualenv 是一个用于创建 Python 虚拟环境的工具, 可以使得多个 Python 之间互不影响, 相互独立。例如: 在应用 A 创建的虚拟环境中, 若某个套件已经在应用 B 中安装, 仍可安装该套件的不同版本, 并且可独立升级, 不影响其他应用中的该套件。

#### c. 数据库配置

为了使得基础框架能够运行出图像界面, 需要对数据库进行配置, 首先需要添加 PostgreSQL DB 用户, 通过该 DB 用户创建数据库, 最后填充数据库表。

其中关键命令如下：

```
$ sudo su postgres
$ createuser -d -P pybossa
$ createdb pybossa -O pybossa
$ exit
$ python cli.py db_create
```

### 2.3.2 “众研”平台的搭建与介绍

“众研”平台构建在 Pybossa 框架之上(一个开源的众包框架)，由于它只支持英语和西班牙语，为了支持国内用户的使用，首先我们扩展了中文语言包，以便平台语言汉化，其次，为了能够支持一些图像标注项目的交互方法研究，所以本小结对其进行二次开发的部分工作(其他将在第三章和第四章介绍)，使其可以支持我们的一些图像标注项目的发布和管理。

#### a. 中文汉化过程

默认情况下，Pybossa 框架只支持两种不同的语言：英语和西班牙语，在使用过程中可以对平台语言进行切换，按照下述步骤可以将中文语言包扩展进去，使其也支持中文。

- (1) 创建翻译文件，执行完后会在当前文件夹创建一个新的 translations 文件夹，里面包含 zh 文件夹

```
$ pybabel init -i messages.pot -d translations -l zh
```

- (2) 使用 Poedit 编辑器打开目录/translations/zh/LC\_MESSAGES/下的 messages.po 文件，翻译成中文。

- (3) 使用下述命令编译该翻译文件。

```
$ pybabel compile -d translations
```

- (4) 更新现有的翻译文件。因为 Pybossa 框架可能会有新的字符串要进行翻译，为了及时更新，可以提取应翻译的新字符串，然后更新现有的翻译文件，然后重新翻译文件即可，主要命令如下：

```
$ cd pybossa
$ pybabel extract . -F babel.cfg -k lazy_gettext -o translations/messages.pot
```

```
$ pybabel update -i translations/messages.pot -d translations
$ pybabel compile -d translations
```

(5) 打开配置文件 `local_settings.py`，然后定位到数组 `LOCALES` 定义行，将 `zh` 参数添加进入，则成功添加了中文语言包，用户登录“众研”平台后，可以在语言设置项中，设置成中文，平台语言即可切换成中文。

## b. 平台介绍

图 2.11 展示了“众研”平台的一些基础功能(主页, 创建项目, 设置项目参数, 查看任务结果, 导出结果等)。雇主可以在平台上创建项目, 在创建时可以为任务添加详细描述, 并设置是否允许匿名用户参与, 任务优先级、任务冗余度等相关属性。项目创建后通过模板导入不同类型的任务, 包括声音模式识别、图像模式识别、文件抄写、视频模式识别等任务。对于每个项目, 任务创建者可以创建博客供工人进行交流。项目导入任务并发布后, 工人可以开始执行项目中的每个任务, 执行过程中, 任务创建者可以通过任务浏览功能在线观看每个任务的执行情况; 同时, 任务创建者可以随时通过结果导出功能将该项目中执行任务的结果导出到本地进行查看。

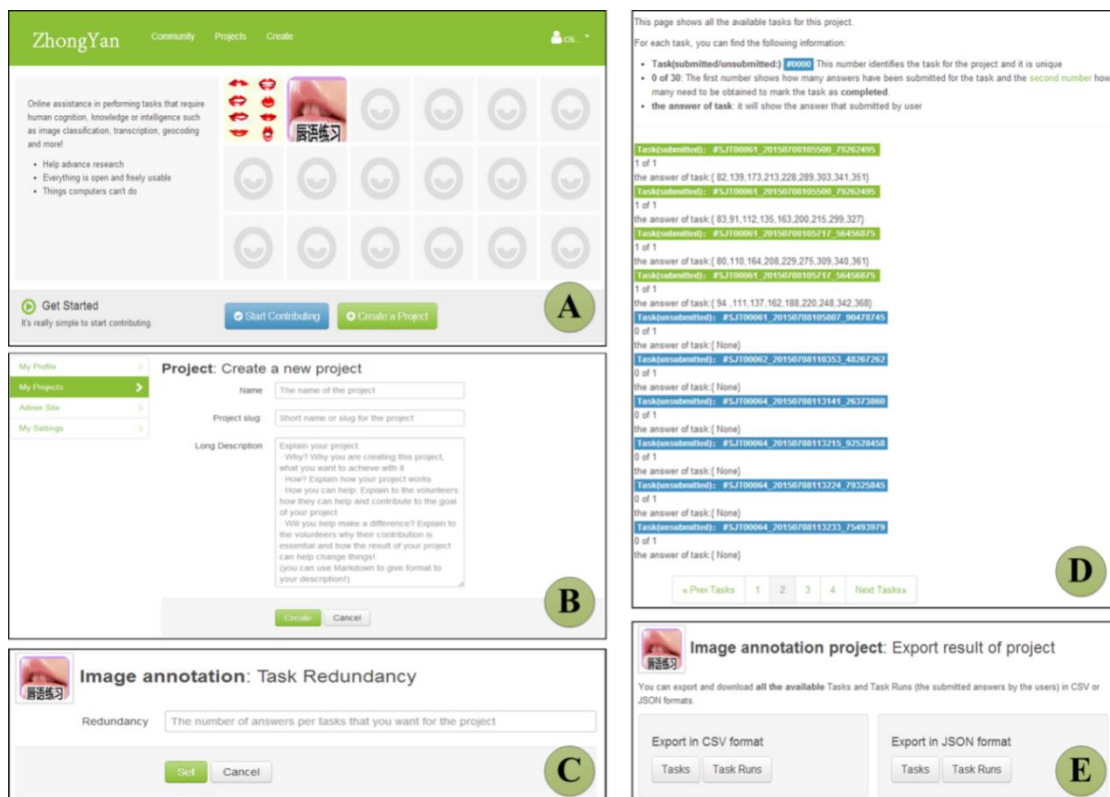


图 2.11 众研平台的部分功能页面展示, (A)主页, (B)创建项目界面, (C)项目设置页面, (D)任务结果展示页面, (E)结果导出页面



## 2.4 本章小结

本章通过对人们日常应用领域, 机器学习以及图像识别等科研领域对于图像标注需求的调研分析, 通过抽象归纳出这些图像标注任务的特性, 大致将其分为了以下几类: 语义标注, 点位标注, 区域标注, 序列筛选标注等。然后简单介绍了我们搭建的众包实验平台-“众研”; 综上所述, 完成图像标注任务分类和平台搭建工作后, 为后续一系列工作提供了相应的研究基础。



## 第三章 众包工作模式的对比性研究

### 3.1 概述

经我们的调研发现[22]，目前完成图像标注任务的工作模式主要有两种：一种是较为传统的线下工作模式，即任务需求方通过一些即时通讯软件联系众包工作者并传输任务文件，主要通过一些图像标注或图像浏览软件在线下完成，然后通过通讯软件将结果提交给任务需求方；另一种是通过众包平台的线上工作模式，也即众包平台作为桥梁，将任务发布方和众包工作者连接在一起，直接在线完成图像标注工作[3]。下面将分别介绍这两种工作模式的基本工作流程以及本章对比性研究的目的。

#### 3.1.1 线下众包工作模式介绍

任务需求方通过一些即时通讯软件联系众包工作者，并传输标注所需的数据给众包工作者，然后提供标注文档，有些复杂的标注任务还需提供配套的图像标注软件与标注文档，众包工作者学习标注过程后，首先需要通过所需的标注软件对接受的图像进行标注，并且在标注的过程中将标注结果按要求记录下来；最后再通过通讯软件将存放标注结果的文件发送给任务需求方。

#### 3.1.2 线上众包工作模式介绍

任务需求方将任务上传并发布到众包平台上，然后众包工作者在众包平台上完成这些图像标注任务，同时，每个任务的标注结果将会存储到众包平台的数据库；任务需求方可以在众包平台上查看任务完成进度，以及浏览任务结果并可导出所有的结果数据。

#### 3.1.3 研究目的

虽然众包提供了一种经济有效的生产方式，但是如上所述，很多类型的众包任务是通过线下众包工作模式来完成，所以相比较而言，通过线上众包工作模式来解决图像标注任务的优缺点在哪？针对图像标注任务的交互方法又该如何设计？为了回答这些问题，我们进行了一个对比研究，同时通过这两种方式来完成同一批图像标注任务，通过定性和定量分析，总结得出一些发现和结论，作为后续交互方法的设计原理启示。

## 3.2 项目和任务介绍

### 3.2.1 项目背景及介绍

随着现代计算机学科的发展,机器学习(Machine Learning),作为人工智能(Artificial Intelligence)的核心,已经渗透到生活中的各个领域,例如机器人、医疗、交通和饮食等[42]。许多研究工作通过机器学习方法来尝试解决图像识别问题,例如使用深度神经网络训练法对图像进行标注[43]。需要注意的是,提前获取足够的标注数据作为训练集是改进机器学习算法和模型的基础。本次在我们对比性研究里所采用的图像标注项目就是在这样一个背景下诞生的。

众所周知,在日常生活中,人与计算机之间常见的交互方式包括:鼠标键盘,语音,手势等;当然还有一些其他的交互方式仍在进一步研究和完善中,例如:眼控,唇语等。其中唇语指的是通过用户的嘴唇变化和状态的一系列图像来确定用户的输出信息,包括文字,数字等。某图像研究机构想通过机器学习的方法来研究用户在读取不同数字时嘴唇变化和状态的对应关系,所以利用唇语标注项目共获取足够多的图像标注数据作为训练集。

按照第二章节中对图像标注任务的分类,唇语标注项目中的任务属于序列筛选标注任务类型,唇语标注项目的基本信息如表 3.1 所示:

表 3.1 唇语标注项目基本信息

项目名称	所属类型	任务数量	图片数量	文件总大小
唇语标注	序列筛选标注	3149	26000	16.2G

### 3.2.2 任务介绍

每个图片文件夹包含一个图片数字序列和上百张图片,是一个人读 8 位验证码的过程。对于这个数字序列中的 8 个数字,需要标注读者在读每个数字的起止位置帧号。

#### 标注过程说明

下面举例说明,以任务序列号 V11\_20150616111243\_85053092 为例,具体标注步骤如下:

- (1) 打开文件夹“V11\_20150616111243\_85053092”,共计 236 张图片,其内容是一个人读数字“85053092”。使用图片查看工具打开第 1 张图片,用键盘中的

左右键开始快速浏览，确定读数“85053092”大致的起始位置。

- (2) 标注起止帧号：用左右键开始慢速浏览，此时可以看到嘴唇开始读第一个数字“8”时对应的是第 97 帧图片，将它记录下来。
- (3) 确定下一个数字的开始位置：继续慢速浏览，确定“5”的起始位置是第 107 帧并将它记录下来，它同时也是上一个数字“8”的结束位置。
- (4) 以此类推，依次确定“0”、“5”、“3”、“0”、“9”、“2”的起始位置并记录之
- (5) 最后确定末位读数“2”的结束位置帧号 172 并把它记录下来。
- (6) 8 个计数总共需要记录 9 个点，将文件夹名以及这 9 个点的位置帧号记录在帧号记录文本文件中，文件夹名和帧号之间以及各帧号之间用一个或多个英文空格隔开。

### 标注结果文档说明

为了便于统计每个任务的标注结果以及众包工作者的任务工作量，任务需求方对标注结果文档的要求如表 3.2 所示：

表 3.2 标注结果信息示例

标注结果信息											
任务文件夹名	帧数									众包工作者	备注
V11_20150616111243_85053092	97	109	119	128	137	144	158	165	172	User_1	
V11_2015061203212_72572572	77	94	104	113	124	132	142	150	159	User_2	

### 检查标准

#### a. 程序自动检查部分

- (1) 每个图片序列需标注 9 个边界，如果一个序列标注的边界数小于 9，则算作 9 个标注错误。
- (2) 每个图片序列的 9 个边界数值需逐个递增，否则算作 9 个标注错误。

#### b. 人工检查部分

- (1) 每个边界的偏差如果超过 3 帧，则算作一个标注错误，如果偏差超过 6 帧，则算作 9 个标注错误。
- (2) 每批随机抽查视频的 5%，确定本批次的标注准确率。

## 3.3 对比性研究方法

为了分别通过两种不同的众包工作模式来完成唇语标注项目，本文招募了

20 名众包工作者(来自残疾人群体)来参与完成这些任务,他们来自不同的地区,并且性别和年龄也不同,如表 3.3 所示。随后我们将这些任务通过这两种工作模式来完成(为了能使得“众研”平台能支持唇语标注项目,本文对图像序列化筛选类型的标注任务进行了相应的设计和实现,具体将在第四章介绍),章节 3.3.1 将介绍通过这两种众包工作模式来完成唇语标注项目的具体过程。

表 3.3 众包工作者的基本信息

众包工作者编号	年龄	性别	地区
No.1	48	男	湖南 湘潭
No.2	59	女	上海 虹口
No.3	47	男	安徽 六安
No.4	45	女	湖北 武汉
No.5	51	女	宁夏 中卫
No.6	49	女	黑龙江 大庆市
No.7	54	男	黑龙江 鹤岗市
No.8	25	男	广东 佛山
No.9	52	男	山东 沾化县
No.10	30	男	江苏 泰州
No.11	21	女	山东 沾化县
No.12	37	男	湖北 武汉市
No.13	47	男	重庆 万州区
No.14	31	女	广东 番禺市
No.15	36	女	广东 从化县
No.16	31	女	广东 韶关市
No.17	26	男	上海 浦东
No.18	32	男	广东 韶关市
No.19	33	女	广东 珠海市
No.20	35	女	广东 珠海市

### 3.3.1 两种众包工作模式下的工作过程

#### a. 通过线下众包工作模式来完成唇语标注任务的步骤

- (1) 用户通过 QQ 或者邮件的方式从任务请求者那里获取完成任务所需的图片资料, 包含 n 个图片文件夹以及对应的数字序列。

- (2) 通过图片浏览工具依次查看每张图片，标注每个读数在该图片序列中的起止位置帧号，并将结果记录在文件中。
- (3) 将记录的文件按照相应格式发送给任务需求方。

#### **b. 通过线上众包工作模式来完成唇语标注任务的步骤**

- (1) 任务需求方在众研平台上创建唇语标注项目，导入需要完成的任务。
- (2) 用户登陆平台，查找相应的任务，点击进入，便可开始做任务。
- (3) 通过翻页按钮，同时查看多张图片，标注每个读数在该图片序列中的起止位置帧号，将结果记录在网页上的结果提交框中。
- (4) 点击提交按钮，即可将结果保存。任务需求方可以通过平台的结果导出功能查看所有用户提交的结果。

### **3.4 数据采集与分析**

首先，介绍我们用来进行本次对比研究的唇语项目介绍：一方面，通过 QQ 发布了 127 组唇语标注任务给部分众包工作者，让他们通过手动方式(线下工作)来完成这项工作，最后通过 QQ 收集了所有众包工作者记录的结果文档；另一方面，在众研平台上导入了任务需求方提供的 3146 组唇语标注任务数据供多个用户线上进行完成。共有 20 位众包工作者用户来完成这批唇语标注项目，在两个星期左右的时间内，所有的任务均已完成，并最终通过平台的任务结果导出功能，任务需求方获得了这 3146 组唇语标注任务的结果。

#### **3.4.1 数据采集**

在这些任务执行和完成之后，我们持续对相关数据进行了采集，这些数据包括“众研”平台日志数据、用户问卷调查及访谈等。具体如下：

##### **a. 日志数据**

从“众研”平台的数据库中，本文收集了以下日志数据，包括每个任务 ID、任务状态、起始时间、标注结果、用户 ID、访问 IP 等。

##### **b. 问卷调查和访谈**

显然日志数据只能统计部分线上工作模式的数据，因此在任务完成之后，我们还设计了一套问卷调查并对一些用户在通讯软件上进行了相关的访谈，以方便获取更多的数据。其中问题主要是围绕以下三个方面来设计的：

- (1) 第一方面主要围绕用户的背景信息，如：用户的性别，年龄，地区，从事互联网工作年限，每日工作时长等。
- (2) 第二方面主要围绕唇语标注项目在两种工作模式下的工作信息，如：任务量及工作时长，工作模式评分等。
- (3) 第三方面主要围绕用户对两种工作模式的主观感受：如：两者优点和缺点的概括，自己更偏爱工作方式，选择的原因等。

### 3.4.2 定量和定性相结合的结果分析

通过将以上采集的数据汇总后，可以计算出所有用户的年龄，本次参与唇语标注项目的众包工作者用户的年龄段分布在[20,60)区间，10.5%属于[20,30)区间,42.1%属于 [30,40)区间,26.3%属于 [40,50)区间, 21.1%属于[50,60)区间。我们对平台的众包工作者用户进行了问卷调查和相关访谈，共收回 11 份有效的结果，通过对结果的分析可得出以下基础结论。

在中国，近十年的互联网迅猛发展，使得他们基本都有一定时长的互联网工作经验，少则数月，多则数年。他们会结合自己的生活规律和作息来合理分配完成这些任务的时间。54.5%的用户全天都比较自由，可以选择在一天的任意时刻来进行网络工作；剩下 45.5%的用户，会选择上午或下午或者晚上来进行网络工作。所以在工作时间上，63.6%的用户每天会分配 6 小时以上的时间来进行互联网工作，剩下 36.4%的用户每天工作时间则少于 6 小时。在我们调查的时候，除了参与唇语标注项目，有 45.5%的用户还同时参与了其他多个互联网工作。

对比两种众包工作模式下的相关数据，通过定量和定性结合分析可知，相较于线下工作模式，使用“众研”平台完成唇语标注任务有以下优点：能提高工作效率，减少对本地资源的占用率以及减少传输大量数据的麻烦等；但缺点也很明显，例如：增加了对网络的依赖程度，限制了完成任务的灵活性等。但总体而言，使用线上众包工作模式更受用户喜爱，具体结论如下所述：

#### (1) 众研平台提高了工作效率

在调查的所有用户中，81.8%的用户同时通过两者方式参与完成了唇语标注任务。通过线下的方式他们平均需要 7.7 分钟来完成一个唇语标注任务。而通过众研平台，他们平均只需要 6.8 分钟来完成一个唇语标注任务，每个任务将近节省了 1 分钟的时间，如图 3.1 所示，主要因为众研平台简化了原有的工作流程，用户能够更快捷地处理和提交任务，所以能够节省一定的时间。



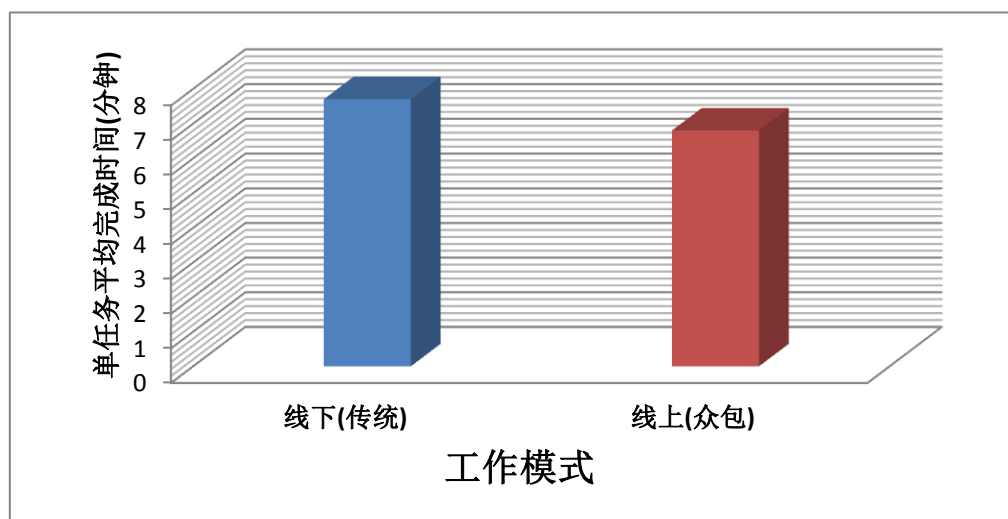


图 3.1 两种模式下用户完成单个任务的平均时长

## (2) 众研平台增加了对网络的依赖程度

对于一批包含多个唇语标注任务的工作，线下众包模式只在两个时间点需要连接网络，分别是在工作开始时通过 QQ 等方式接收任务图片资料；以及任务完成后将任务结果发还给雇主，除这两个时间点之外，即便没有网络连接，用户也能在自己的电脑上继续完成任务。

而通过众研平台，则在每个任务的浏览和结果提交上都需要保持网络的通畅，详细的说，众研平台每个页面的跳转和变化都需要网络的支持，如果某一时刻网络不能联通，当前则不能顺利的完成正在进行的标注任务。对比这两种方式，我们可以发现，线下众包模式对网络的依赖程度较低。

## (3) 众研平台减少了对本地资源的占用率

首先，线下众包模式需要占用自己个人电脑的硬盘空间，虽然一个唇语标注任务的图片平均只有 5MB 的大小，但是如果用户一次性接收大量任务，那么将占用一定的个人电脑的硬盘空间。例如：3146 个唇语标注任务的文件大小为 18.3GB。

然而通过众研平台，每次通过浏览器进行页面访问，只会造成极小的页面缓存。对比这两种情况可以得出，通过众研平台完成任务对个人电脑的资源占用更低。

## (4) 众研平台的灵活性偏弱

线下众包模式的灵活性主要体现在任务个性化完成上，就唇语标注任务而言，

用户可以选择使用自己喜欢的图片浏览工具对单个图片进行查看,随意调整图片的大小甚至角度,并且可以控制单个图片浏览的速度,能够更好的比较相邻图片中嘴唇的变化差异。平台在图片切换过程中则需要考虑网速的影响,由于网速波动的影响,页面在加载过程中可能出现卡顿的现象,影响任务完成的流畅性。

众研平台的灵活性主要体现在数据全部存储在服务器云端上,用户可以在任意联网的终端上(PC,平板或手机)通过网页便可以随时随地的查看这些数据,并完成相应任务。而线下众包模式只能在存储数据的个人电脑上完成任务,一旦用户与其个人电脑分开,则无法继续做任务。

### (5) 线上众包工作模式更受用户偏爱

综合比较这两种完成任务的方式,在满分 5 分的前提下,用户给这线下传统的众包工作模式给出了平均分 3.78 分,给众研平台方式给出了平均分 4 分。在偏爱的工作方式调查中,有 73%的用户表示他们更愿意通过众研平台来完成任务,如图 3.2 所示。主要原因是,众研平台能够节省完成任务的时间,提高了工作效率,可以争取更多的报酬。另外但是有 27.2%的用户表示他们仍然喜欢原来的手动方式,主要原因是,手动方式不受网络限制,图片浏览形式可以自由选择。

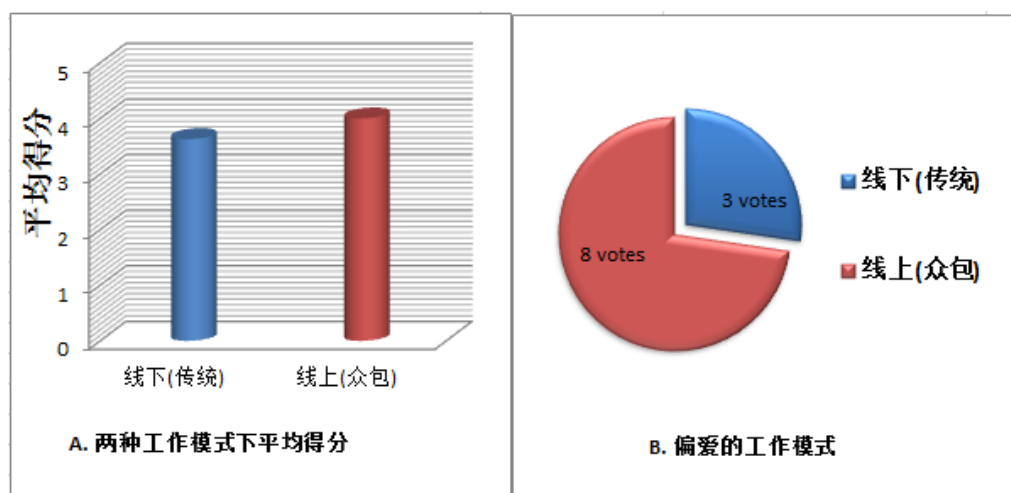


图 3.2 两种工作模式的平均得分和得票统计

## 3.5 交互方法设计原理启示

本章节发现,在当前众包工作模式环境完成图像标注任务的过程中,众包工作者们基本上把工作效率放在第一位,因为参与网络兼职的他们,希望能够在同等时间内,获取更多的报酬;其次受限于他们的年龄和教育层次,并不是所有的众包工作者对于网络工具的操作都十分娴熟,所以他们希望在完成图像标注任务

的标注方式上能够更加简便和自由，避免给他们带来一些认知上的操作障碍，使得他们能够很轻松的上手并胜任这些任务。基于此发现，我们提出了在众包平台上针对图像标注任务的交互方法在设计上应该特别注意的几个地方：

- (1) 首先，在设计交互方法时，应该把图像标注效率放在第一位，促使用户能够在众包平台上更快速的完成标注。
- (2) 其次，考虑到众包工作者的年龄和教育层次不同，所以应当尽量将图像标注任务的交互过程设计的尽量简便，减小众包工作者与系统之间的认知障碍。
- (3) 最后，在每个任务提交过程中，若该任务类型对结果有基础的逻辑限制，应当在任务结果提交时添加逻辑验证模块，帮助众包工作者提前规避掉一些基础的逻辑错误，从而提高标注准确度。

### 3.6 本章小结

本章首先通过介绍当前完成图像标注任务的两种工作模式，一种是较为传统的线下工作模式，另一种是通过众包平台的线上工作模式；随后进行了一个比较性研究：从结果上看，这两种工作模式都能完成图像标注任务，但是他们的工作方式有许多不同之处，例如：任务数据的传输过程，任务标注过程，任务结果获取过程等。那么我们在设计交互方案的时候该采取什么样的依据？为此，我们对这两种众包工作模式进行了对比性研究，试图去了解这两种工作模式在完成图像标注任务时的优缺点，通过搭建的众包实验平台“众研”对两种工作模式进行了对比分析并得出一些结论，将此作为我们设计图像标注任务交互方法的原理启示。



## 第四章 图像标注任务的交互方法设计与实现

### 4.1 概述

本章将主要面向线上众包平台中的图像标注任务提出相应的交互方案，并在“众研”实验平台中介绍相关的设计和实现过程。

首先，本文在前两章中对图像标注任务进行了调研，通过对任务之间的共性抽象，将当前众包工作中所面临的主要图像标注任务分为了四类。其次，搭建了“众研”实验平台，通过一些基础的二次开发，使其能基本支持序列化筛选类别的图像标注任务。然后，对两种众包工作模式进行了对比性研究，通过数据分析得出了一些交互方法的设计启示。

因为实际的图像标注任务比目前众包平台上的一些基础性的标注任务要复杂很多，所以本章将在这些基础上，首先针对两类用户的实际需求进行调研分析，并结合第三章的设计启示总结出设计准则，然后基于“众研”实验平台，对主要面临的四类图像标注任务的交互方法进行方案设计、实现和优化，同时将对设计和实现过程中的原理，关键技术和算法进行介绍。

### 4.2 需求调研与问题分析

首先介绍当前“众研”平台中，任务上传以及任务标注两个过程的操作说明，在之间的对比性研究的基础之上，本章节为了使“众研”平台能够更好的支持图像标注任务，本文通过问卷、微信、QQ 语音或视频聊天等方式对其中 11 名众包工作者以及 2 名任务需求方进行了调研，如表 4.1 所示，分析用户在“众研”平台上针对这两个操作过程的理解和认识，主要询问的问题包括以下几个方面：

- (1) 第一方面主要针对任务需求方在任务上传过程中涉及的相关问题，包括：自身计算机知识背景，整理上传文档的耗时，上传文档过程中遇到的麻烦，经验总结以及改进建议等。
- (2) 第二方面主要针对众包工作者在任务标注过程中涉及的相关问题，包括：工作使用的设备，标注难点以及工作中自身情绪波动，和经验总结等。

下面通过任务上传和标注过程介绍，结合调研总结，归纳出对应过程中待解决问题。

表 4.1 调研对象基本信息

用户编号	年龄	性别	地区	用户身份
No.1	35	女	浙江 杭州	任务需求方
No.2	26	男	上海 浦东	任务需求方
No.3	48	男	湖南 湘潭	众包工作者
No.4	59	女	上海 虹口	众包工作者
No.5	47	男	安徽 六安	众包工作者
No.6	45	女	湖北 武汉	众包工作者
No.7	51	女	宁夏 中卫	众包工作者
No.8	49	女	黑龙江 大庆市	众包工作者
No.9	54	男	黑龙江 鹤岗市	众包工作者
No.10	25	男	广东 佛山	众包工作者
No.11	21	女	山东 沾化县	众包工作者
No.12	37	男	湖北 武汉市	众包工作者
No.13	47	男	重庆 万州区	众包工作者

#### 4.2.1 任务上传过程介绍

任务需求方想在“众研”平台上创建项目，并导入一至多个任务的操作步骤如下：首先需要在平台上创建项目，将任务中的图片文件，部署到可访问的图片服务器，确保每张图片有一个可访问的 URL，汇总这些 URL，然后将其输入到谷歌电子表格中，生成链接，通过链接导入任务的形式，在这些任务上传到项目中。

#### 待解决问题

- (1) 首先，对于任务需求方而言，需要自己提前将图片文件部署到可访问的图片服务器，这对于许多非计算机专业人士而言比较麻烦。
- (2) 第二，整理汇总图片 URL 比较麻烦，如果遇到任务量较大的情况，则会消耗较长的时间去整理任务上传文档。
- (3) 第三，受限于 Pybossa 框架的影响，只能通过谷歌电子表格，DropBox 等国外文档处理网站的链接上传任务文档，国内用户需要购买 VPN，才能够成功访问这些网站和生成链接，然后将此文档链接导入项目任务中。
- (4) 第四，任务发布后，有时需要调整标注界面和交互方法时，任务需求方无法直接修改相应的标注界面，需要平台管理员在后台帮助修改。

### 4.2.2 任务标注过程介绍

众包工作者首先需要通过输入账号和密码，登录到“众研”平台中；通过选择相应的项目，点击做任务从而进入图像任务标注界面。对于不同类型的图像标注任务，系统会适配不同的界面风格和交互方法；以图像语义标注任务界面为例，通过两栏的形式将问题以及可选答案放在左边显示，相应图像在右侧显示。众包工作者通过浏览问题以及图片，最后在可选答案框里选择相应的答案点击鼠标左键，即可完成该任务，随后转至下一个任务的标注界面。

#### 待解决问题

- (1) 首先，对于众包工作者而言，他们工作在不同的终端，图像显示尺寸会有所变化，其次有时需要将图片放大或者缩小以便更好的完成标注工作。
- (2) 第二，有些标注过程容易出错，且难以修改，将导致要重新标注一遍，比较费时。
- (3) 第三，相近任务重复较多时，容易乏味，想自主切换到其他任务界面。
- (4) 第四，标注过程比较繁琐，导致标注疲劳，并影响工作效率。

## 4.3 设计原理与技术方案

### 4.3.1 设计理论

#### 用户参与式设计

用户参与式设计(User Participatory design)概念起源于 60 年代北欧国家，强调让用户的声加入设计决策制定过程中[44]。这一设计理论目前在很多领域中被广泛使用，主要包括城市设计、工具设计、软件或产品开发等。

### 4.3.2 设计准则和方案

基于用户参与式设计理论，本文在第三章总结了使用众包平台的几个优缺点，例如：可以提高工作效率，减轻对本地资源的占用率，避免多用户之间文件大数据的传输等。同时，使用众包平台也带来了一些缺点，例如：使得完成标注任务的过程十分依赖于网络的联通性，没有提供个性化的任务完成方式等，并总结了针对图像标注任务交互方法的设计原理启示。其次，在前两小节中，总结了当前“众研”平台中待优化和解决的一些问题。

综合以上,为了能够更好的在“众研”平台上支持图像标注任务的发布与标注交互过程。本文总结了交互方法的一系列设计准则和方案:

- (1) 搭建图片服务器,编写任务上传文件自动生成脚本,辅助任务需求方更快速生成任务导入文件。
- (2) 实现项目任务可通过本地 CSV 文件导入,丰富导入任务形式,辅助任务需求方更方便的完成任务上传发布工作。
- (3) 针对已经具体分类的四个类别的图像标注任务进行具体的方案设计,构建四类交互模板,并且使其可视化,可由任务需求方自己动态更新。
- (4) 考虑到众包工作者来自社会不同阶层,任务交互界面应尽量简洁易懂,能够快速上手。
- (5) 简化任务的标注过程,为了提高工作效率,使得众包工作者能够在同等时间完成更多的众包任务,要尽可能赚取报酬。
- (6) 为一些对结果有基础限制的任务,添加结果检验机制,提高用户标注准确度。
- (7) 尽可能的为众包工作者提供多样化的完成和标注方式选择,从而提高用户体验。

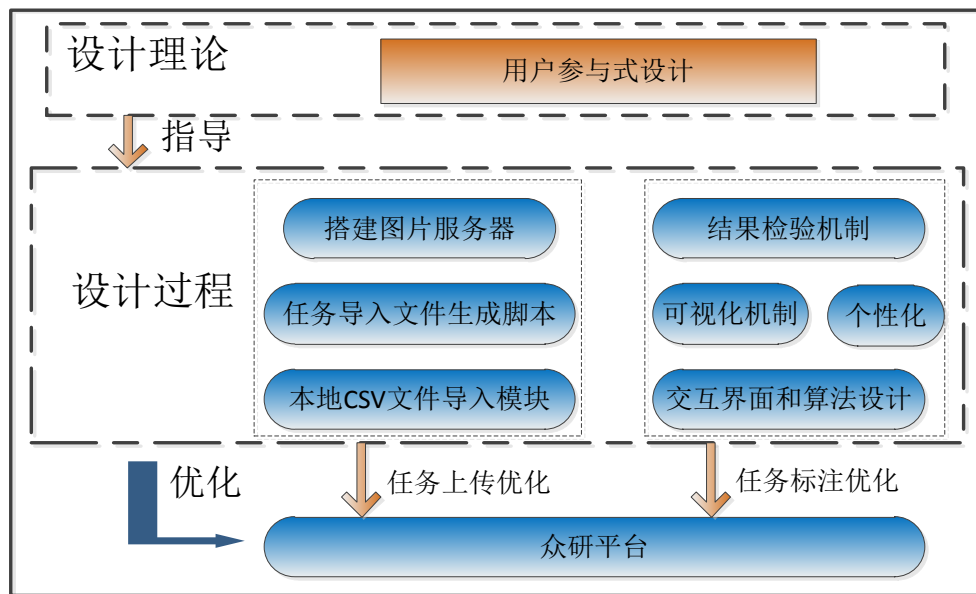


图 4.1 交互方法设计原理图

### 4.3.3 交互模板方案设计概要

下面按照对图像标注任务的四种分类,分别介绍对应的交互模板方案设计。



### (1) 语义标注

- **界面布局:** 采用两栏的形式, 左边栏要显示的内容包括: 问题描述, 可选答案, 任务进度条, 提示框等; 右边栏要显示的内容为: 图片框(要来加载待标注的图片)。
- **内容来源:** 其中待标注图像的 URL、问题、可选答案以及提示内容来自于任务发布方; 任务进度以及当前任务 ID 来自于数据库查询。
- **标注的方式:** 众包工作者通过鼠标左键点击可选答案。
- **记录标注结果的方式:** 可选答案框触发点击事件则记录该任务的标注结果。
- **提交结果并进入下一个任务的方式:** 众包工作者点击可选答案框, 系统记录标注结果后, 跳转至下一个任务的标注界面。

### (2) 点位标注

- **界面布局:** 采用多行多列的形式, 第一行采用单栏的形式, 用来显示提示框与提示内容; 第二行采用单栏的形式, 用来显示任务进度条以及任务 ID, 第三行采用三栏的形式, 第一栏要显示的内容为图片框(要来加载待标注的图片), 占该行宽度比 50%, 第二栏要显示的内容为可拖拽的带序号的图标, 占该行的宽度比 10%。第三栏要显示的内容为一个表格, 包含序号, 点位描述, 点位坐标。
- **内容来源:** 其中待标注图像的 URL、标注点位框答案以及提示内容来自于任务发布方; 任务进度以及当前任务 ID 来自于数据库查询。
- **标注的方式:** 众包工作者通过鼠标左键拖拽标注点图片到图片框的相应点位上完成标注。
- **记录标注结果的方式:** 左键选中某个拖拽图标时, 进行拖拽动作将其拖拽至图像中某个部位松开时, 系统需要记录该点位在图像中的像素坐标点, 并将其在答案表格框中, 动态显示。
- **提交结果并进入下一个任务的方式:** 点击提交按钮时, 系统记录答案后, 跳转至下一任务的标注界面。

### (3) 区域标注

- **界面布局:** 采用多行多列的形式, 第一行采用单栏的形式, 用来显示提示框与提示内容; 第二行采用单栏的形式, 用来显示任务进度条以及任务 ID, 第三行采用两栏的形式, 第一栏要显示的内容为可选颜色列表, 第一栏要显示的内容为可选形状列表, 第四行采用两栏的形式, 第一栏为图片框(加载带标注的图片), 第二栏要显示的内容为一个表格, 包含序号, 区域描述, 中心点

坐标, 半径或边长。

- **内容来源:** 其中待标注图像的 URL、标注区域颜色、形状以及提示内容来自于任务发布方; 任务进度以及当前任务 ID 来自于数据库查询。
- **标注的方式:** 众包工作者可选择标注颜色以及标注形状(不选的话会使用系统默认的颜色和形状), 通过鼠标左键直接在图片的相应区域上点击绘制该形状完成标注。
- **记录标注结果的方式:** 绘制完标注形状后, 系统需要记录该点位在图像中的中心坐标点以及半径或边长, 并将其在答案表格框中, 动态显示。
- **提交结果并进入下一个任务的方式:** 点击提交按钮时, 系统记录答案后, 跳转至下一任务的标注界面。

#### (4) 序列筛选标注

- **界面布局:** 采用多行多列的形式, 第一行采用单栏的形式, 用来显示提示框与提示内容; 第二行采用单栏的形式, 用来显示任务进度条以及任务 ID; 第三行采用单栏的形式, 交叉显示待标注结果和序列号; 第四行采用多栏的形式, 动态加载显示多张图片。
- **内容来源:** 其中待标注图像的 URL、序列号以及提示内容来自于任务发布方; 任务进度以及当前任务 ID 来自于数据库查询。
- **标注的方式:** 众包工作者通过浏览动态加载出来的图片, 确定序列号, 并将其手动填写至答案记录框, 填完则完成标注。
- **记录标注结果的方式:** 待标注结果框可供众包工作者输入相应的图片序列号, 输入后则记录该答案。
- **提交结果并进入下一个任务的方式:** 点击提交按钮时, 系统记录答案后, 跳转至下一任务的标注界面。

#### 4.3.4 关键技术和算法

在交互方法的实现过程中, 运用了多种技术。例如: 在任务导入文件数据处理过程中, 使用 C++ 语言编写该文件自动生成脚本。“众研”平台在 Pybossa 框架基础上, 使用 Flask 开发框架, 主要采用 MVC 的模式下通过 Python 语言进行开发, 在前端交互界面中采用 HTML+CSS+jQuery+JavaScript 的方式。其中举例介绍在图像点位标注和图像区域标注两个任务模板中主要采用的交互流程和部分算法, 如下所示:

## 交互流程和算法

为了实现用户与标注界面和系统之间的交互，设计了相应的交互流程和算法，其中以图像点位标注任务为例，做相应的算法控制介绍，具体交互流程如图 4.2 所示：

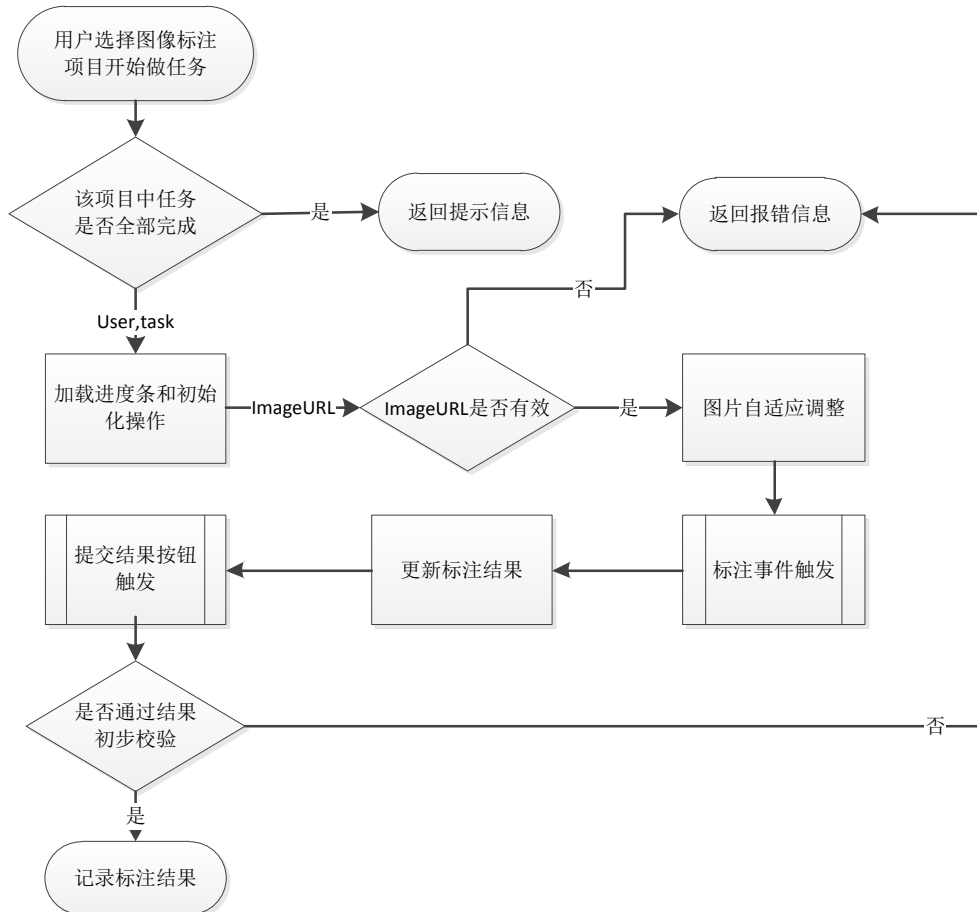


图 4.2 交互流程图

首先，当用户点击做任务按钮后，进入图像标注页面时将会触发任务显示函数，此时，当前任务和标注用户将作为系统的输入参数，根据输入参数，首先需要验证该任务是否全部完成，若未完全完成，则提示相关的提示信息，否则需要加载用户任务完成进度条，初始化标注点位和答案记录表格。然后验证图片 URL 的有效性，根据图片原始大小和用户界面尺寸作相应的自适应调节。顺利的话，根据标注点位移动触发事件，记录点位像素数组并将其更新到前台答案记录表格中显示。最后若用户点击提交结果按钮，则可根据提前定义的结果规则去验证结果的有效性。若有效，则返回记录答案。否则返回相关报错信息，部分关键函数具体的算法实现将在算法 4.1-4.4 中给出。

在上述交互流程的指导下，算法 4.1 描述了整体交互算法实现的大致过程，

整体算法过程用伪代码形式展现。

---

#### 算法 4.1: 图像标注交互算法

---

输入: *user*: 用户, *task*: 当前任务

输出: *answer/message*: 标注结果或其他信息

算法过程:

1. ImageAnnotationTask(*user*, *task*) do
  2.   If(!isEmptyObject(*task*)) do
  3.     LoadUserProgress(*user*)
  4.     InitSomeVariables()
  5.      $imageUrl \leftarrow task.info.imageUrl$
  6.     If(CheckPicurl(*imageUrl*)) do
  7.       DrawImage(*ImgD*, *FitWidth*, *FitHeight*)
  8.     Else
  9.       return *WrongMessage*
  10.   Else
  11.     return *RemindMessage*
  12.   End if
  13.   While TriggerEvent() do
  14.     For each  $pos \in posArray$  do
  15.        $pos.coordinate \leftarrow pos.currentCoordinate$
  16.     End for
  17.   End while
  18.   If AnswerSubmit() then
  19.      $answer \leftarrow SetAnswer(posArray, user)$
  20.     If CheckAnswer(*answer*) then
  21.       return *answer*
  22.     Else
  23.       Return *WrongMessage*
-

---

24. End if

25. End

---

任务方上传图片文件到图片服务器后, 首先需要对这些图片的 URL(一般由 IP 地址+端口+图片存储的相对路径构成)进行汇总, 根据图像标注任务的类型 TaskType(因为不同类型的图像标注任务, 所需的列数和列名不同, 也即对应的上传文档格式不同), 整理成 CSV 文档, 然后一键导入到相应的项目, 从而完成图像任务上传, 算法 4.2 将描述“众研”平台通过算法自动化生成任务上传文档的过程。

---

#### 算法 4.2: 任务上传文档自动化生成算法

---

输入: *ServerIp*: 服务器地址, *port*: 端口号, *FilePath*: 图片相对路径, *TaskType*: 任务类型等

输出: *UploadFile*: 任务上传文档

算法过程:

1. FileAutomateGenerate(*ServerIp*, *port*, *FilePath*, *TaskType*) do
  2. For each *image* store in *FilePath* do
  3.     *ImageURL*  $\leftarrow$  GenerateURL(*ServerIp*, *port*, *FilePath*, *image.name*)
  4.     *OtherInformation*  $\leftarrow$  GenerateOtherData(*TaskType*)
  5.     *RowRecord*  $\leftarrow$  GenerateOneRecord(*ImageURL*, *OtherInformation*)
  6.     *AllRecord*[*i*]  $\leftarrow$  *RowRecord*
  7. End for
  8. Create a *UploadFile*
  9. *UploadFile*  $\leftarrow$  WriteToFile(*AllRecord*)
  10. End
- 

在图像点位标注类型的任务中, 要对图像中的某些点位进行比较精确的点位标注, 需要对当前点位在图像中的像素坐标点, 进行记录。因为不同的任务所对应的图片大小不同, 为了让用户能够有一个好的标注体验, 我们首先对图片进行自适应调整, 以一个固定的尺寸展现给用户, 然后在标注过程中, 通过计算像素点的相对坐标和自适应调整比例, 得出一个真实的图像标注点位的像素坐标。算法 4.3 将描述图片自适应调整以及计算点位的实际像素坐标的算法过程。

---

**算法 4.3: 图片自适应调整和标注点位记录算法**


---

**输入:** *Image*: 加载的图片对象, *FitHeight*: 默认展示最大高度, *FitWidth* 默认展示最大长度

**输出:** *PointCoordinate*: 标注点位在图片中的实际像素坐标

**算法过程:**

1. Part 1: 图片自适应调整
  2. ImageAdjustment(*Image*, *FitHeight*, *FitWidth*) do
  3.   *image*  $\leftarrow$  Generate a image object
  4.   *image.src*  $\leftarrow$  *Image.src*
  5.   If *image.width/image.height*  $\geq$  *FitWidth/FitHeight* do
  6.     *Image.width*  $\leftarrow$  *FitWidth*
  7.     *Image.height*  $\leftarrow$  (*image.height* \* *FitWidth*) / *image.width*
  8.     *ScalingRadio*  $\leftarrow$  *image.width* / *FitHeight*
  9.   Else
  10.    *Image.height*  $\leftarrow$  *FitHeight*
  11.    *Image.width*  $\leftarrow$  (*image.width* \* *FitHeight*) / *image.height*
  12.    *ScalingRadio*  $\leftarrow$  *image.height* / *FitHeight*
  13.   return *ScalingRadio*
  14. End
  15. Part 2: 计算点位的实际像素坐标
  16. CalculateRealCoordinate() do
  17.   *RelativeCoordinate*  $\leftarrow$  CalculatePointCoordinate()
  18.   *scalingRadio*  $\leftarrow$  ImageAdjustment(*Image*, *FitHeight*, *FitWidth*)
  19.   *x*  $\leftarrow$  *RelativeCoordinate.x* \* *scalingRadio*
  20.   *y*  $\leftarrow$  *RelativeCoordinate.y* \* *scalingRadio*
  21.   *PointCoordinate*  $\leftarrow$  {*x*, *y*}
  22.   return *PointCoordinate*
  23. End
- 

在部分项目的任务中(例如本文中提到的唇语标注项目中的任务), 任务需求

方根据其自身的标注结果需求和规则 *RulesOfAnswer*，对标注结果 *Answer* 有一定的限制，从而在用户提交结果前去规避一些简单的标注错误，并给出了相关的文档描述，算法 4.4 将描述唇语标注项目中对于结果初步校验的算法过程。

---

**算法 4.4: 结果初步校验算法**

---

**输入:** *RulesOfAnswer*: 标注结果需求和规则, *answer*: 任务结果

**输出:** *true/false*: 是否符合校验规则

**算法过程:**

1. *CheckAnswer(answer, RulesOfAnswer)* do
  2.   For each *rule*  $\in$  *RulesOfAnswer* do
  3.     If *answer* do not conform to *rule* then
  4.       return false
  5.   End if
  6. End for
  7.   return true
  8. End
- 

## 4.4 模块和界面展示

基于以上设计准则和设计方案，我们实现和优化了这些交互功能和界面，下面列出部分模块和界面展示。

### 4.4.1 可视化功能优化

一个项目可以包含多个任务。由于在系统中提供可视化非常有利于系统使用者之间的信息沟通[41]，所以我们为项目的每个任务页面提供了任务 ID 以及当前进度的提示框，其次为了方便用户切换到自己想优先完成的任务，提供了任务导航功能，用户可以根据任务编号直接定位到自己想执行的任务页面中。



图 4.3 任务进度框和任务导航功能

4.4.2 任务生成导入功能优化

原有任务导入模块

相关操作页面如图 4.4 所示，首先任务需求方需要整理汇总图片 URL 比较麻烦，如果遇到任务量较大的情况，则会消耗较长的时间去整理任务上传文档。受限于 Pybossa 框架的影响，只能通过谷歌电子表格，DropBox 等国外文档处理网站的链接上传任务文档，国内用户需要购买 VPN，才能够成功访问这些网站和生成链接，然后将此文档链接导入项目任务中，整个过程比较繁琐。

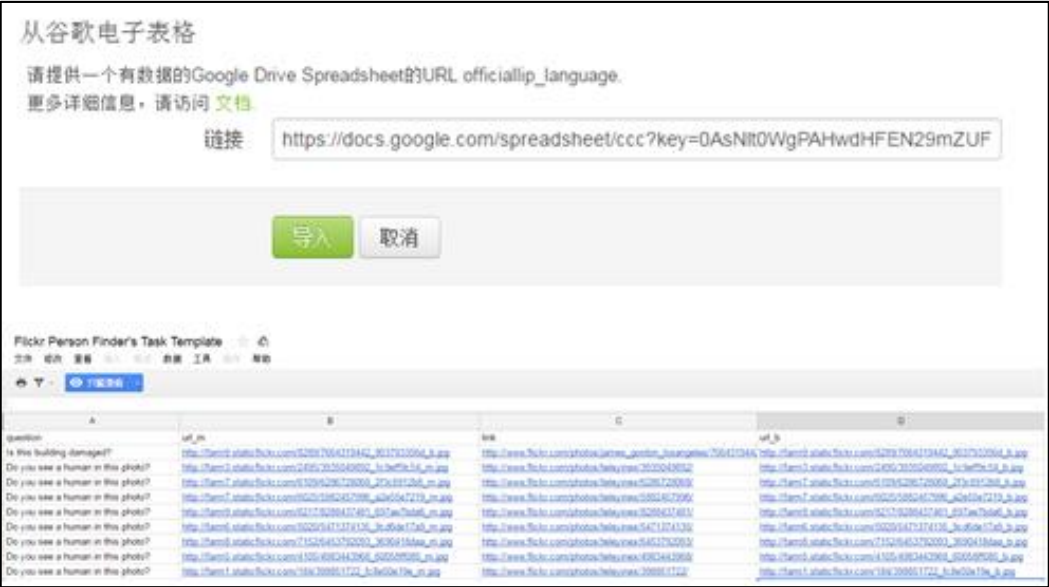


图4.4 以谷歌表格文档的形式导入模板任务



优化后通过脚本自动生成本地 CSV 文件，并设计提供本地 CSV 导入任务功能

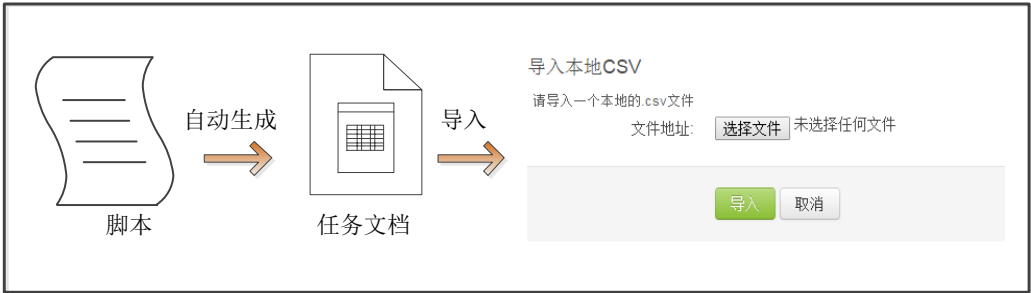


图4.5 优化后导入任务模块

4.4.3 标注交互界面展示

本文准备了一些示例图像数据，将它们上传至“众研”平台服务器，并针对每种类型的图像标注任务，通过给出一个具体实例的方式，来介绍每种交互界面的设计实现过程。

(1) 语义标注交互界面示例

图像语义标注大多通过给出图片并附加问题的方式，要求众包工作者记录答案的形式来完成标注过程。如下给出的具体语义标注任务示例(如图 4.6 所示)是：任务方有一些包含人物表情状态的图片，需要记录每张图片里的人物具体是什么表情，另外，任务方可以提供一些常见的表情(例如：喜悦，愤怒，恐惧，惊喜，疑惑，兴奋等)作为可选项。



图 4.6 图像语义识别任务交互界面

在实现过程中，我们使用图像标签去加载显示待标注的图片，文本标签去加载和显示问题描述，单选按钮用来加载和显示答案的待选项，另外针对一些特殊情况(图片加载未成功或者判断不出是哪种面部表情)，额外准备了两种答案(无图片，不知道)以供标注。众包工作者通过使用鼠标点击这些可选的答案框，后台记录此任务答案，从而完成该任务的标注，并直接跳转至下一个任务的交互界面。

(2) 点位标注交互界面示例

图像点位标注大多通过给出图片并附加待标注点位信息的方式，要求众包工作者记录每个点位的具体坐标形式来完成标注过程。如下给出的具体点位标注任务示例(如图 4.7 所示)是：任务方有一些包含人物全身照的图片，需要记录每张图片里的人物的一些身体部位在图片中的具体坐标位置，任务方需要提供待标注点的数量及描述(例如：点位 1-头部，点位 2-颈部，点位 3-左肩，点位 4-右肩等)。

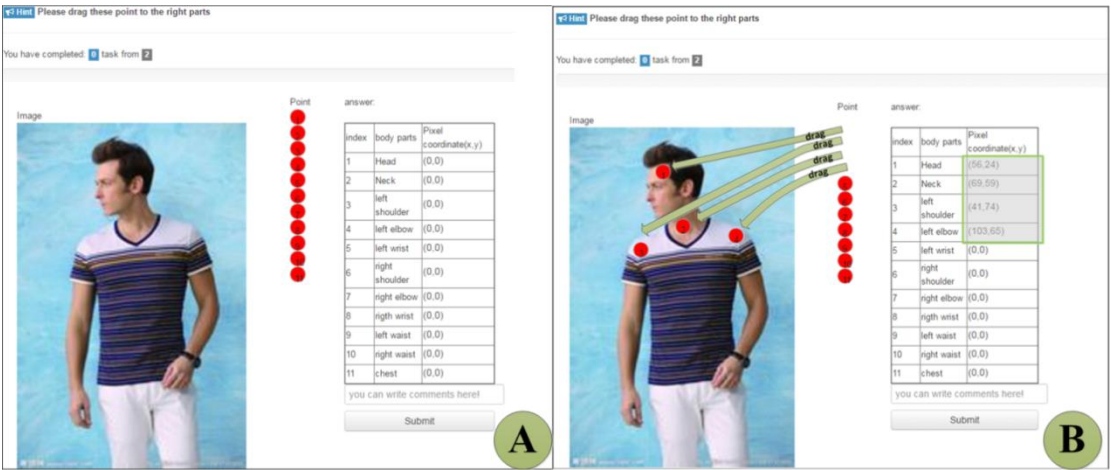


图 4.7 图像点位标注任务交互界面

在实现过程中，我们使用图像标签去加载显示待标注的图片，使用三列多行的表格去加载和显示待标注的点位，分别是序号，部位名称，点位坐标。另外，提供一个文本备注框方便用户记录相关的备注信息。为了简化操作过程和提高灵活性，通过开放标注点图标的可拖拽属性，允许用户直接将相应的点位拖拽至对应的位置处，并更新此点位的坐标值到表格中；每个待标注点位的图片支持多次拖拽，并且也会实时自动地更新点位的坐标值，这种灵活的点位标注方式能够简化操作过程，从而提高了标注效率。最后，标注者通过点击提交按钮，后台记录此任务答案，从而完成该任务的标注，并直接跳转至下一个任务的交互界面。

(3) 区域标注交互界面示例

图像区域标注大多通过给出图片并附加待标注对象的方式，要求众包工作者记录每个对象区域的具体范围的形式来完成标注过程。如下给出的具体区域标注任务示例(如图 4.8 所示)是：任务方有一些包含交通指示标志的图片，需要记录每张图片里某些交通标志的具体区域位置，任务方需要提供待标注的交通标志的说明 (例如:限速标志，交通指示灯标志，斑马线标志等)。

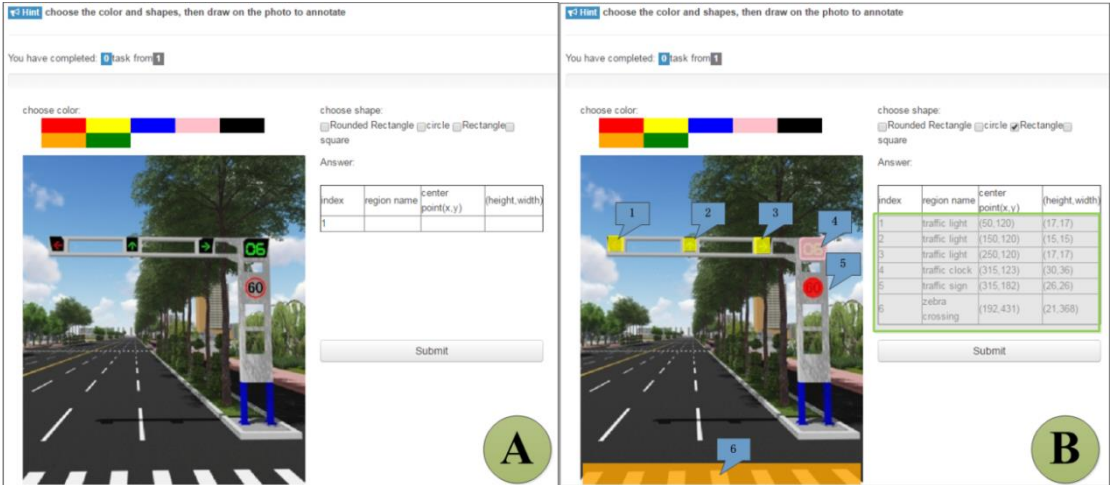


图 4.8 图像区域标注任务交互界面

在实现过程中，我们使用图像标签去加载显示待标注的图片，使用四列多行的表格去加载和显示标注好的对象，分别是序号，对象名称，中心坐标，长和宽。另外，提供一个文本备注框方便用户记录相关的备注信息。首先用户需要选择标注的形状和颜色，然后可直接在图片相应的区域绘制该形状，后台则更新此区域的中心点坐标和长宽值到表格中，这种提供多种形状和颜色的方式，能够适应于多种不同形状的对象，另外对于有交叉覆盖的对象，使用不同的形状和颜色可以更好的区分。最后，标注者通过点击提交按钮，后台记录此任务答案，从而完成该任务的标注，并直接跳转至下一个任务的交互界面。

#### (4) 序列筛选交互界面示例

图像序列筛选标注大多通过给出多张图片并附加指定图片序列描述的方式，要求众包工作者记录某些指定的图片序列位置来完成标注过程。如下给出的具体序列筛选标注任务的示例(如图 4.9 所示)是：任务方有一些包含用户读取验证码的多张连续图片，图片连贯起来能够看清用户在读验证码过程的嘴唇变化趋势，需要记录验证码中每个数字在图片序列中开始和结束的具体位置，任务方需要提供验证码所包含的数字。

图 4.9 图像序列筛选任务交互界面

在实现过程中,我们使用图像标签以多行多列的形式去加载显示多张序列图片,以分页的形式去切换显示不同页码的图片,同时提供个性化设置功能,允许用户去更改每页显示的图片数量,以及图片的大小。另外提供一个文本备注框方便用户记录相关的备注信息。首先用户通过分页浏览去定位到验证码每个数字的起始和结束位置(图片序号),然后将其填写到答案记录框,记录完毕后,标注者通过点击提交按钮,后台记录此任务答案,从而完成该任务的标注,并直接跳转至下一个任务的交互界面。

#### 4.4.4 交互模板功能优化

任务发布后,有时需要调整标注界面和交互方法时,任务需求方无法直接修改相应的标注界面,需要平台管理员在后台帮助修改。经过优化后,本文将此功能修改更新权限同时授予了任务需求方(也即项目创建者),将其可视化展示在页面中,方便任务需求方随时对其进行动态编辑,当其点击更新项目演示器时,则交互界面模板也随即更新,相关页面如图 4.10 所示:



图 4.10 任务模板编辑器

4.5 本章小结

本章首先介绍了用户完成图像标注任务的过程，主要是与任务展示界面的交互；随后通过所使用的具体项目-唇语标注项目对用户进行具体调研，围绕用户为中心得出了对任务交互界面的设计依据，并以此为每种分类进行一对一的交互模板和算法设计，随后针对图像任务上传和标注两个过程介绍了相关模块的优化设计与功能实现。



## 第五章 平台交互方法的评估

本文第二章至第四章中,首先介绍了对图像标注任务的分类以及众包平台的搭建工作,然后分别对图像标注工作中用户上传任务过程以及用户标注任务过程进行了调研,提出了优化后的设计方案并将其实现,但是在实际解决图像标注任务的过程中,是否可以达到较好的效果?比起现有的线上众包平台解决图像标注任务的方案是否能更好的解决在第一章问题领域里面提到相关问题?这些都需要进一步的评估和验证,同时也是衡量系统交互方法可用性的过程,不仅如此,对系统交互方法进行性能评估的意义还包含以下几个方面:

首先,在使用系统交互方法完成图像标注任务的发布和标注能够更好的获取用户的实际需求。其次,通过数据分析能够更加有侧重点的找出交互方法中存在的不足,从而进行改进和优化。

因此本章将通过实验设计,将“众研”平台上提供的图像标注任务解决方案与目前主流的线上众包平台提供的解决方案进行对比评估。

### 5.1 实验设计

本文首先搭建了众包实验平台-“众研”,在“众研”的基础上设计并优化了一种针对图像标注任务的交互方法,试图去解决当前问题领域中所面临的一系列问题。为了更加全面和系统的评估该交互方法的有效性,本文将通过多个实际的图像标注项目,将其与目前市场上众包平台的解决方案进行对比。从而去验证本文提出的交互方法是否比当前的许多众包平台上所提供的方案能更好的支持用户解决图像标注任务。

在第二章至第四章中,本文首先将图像标注任务分为了四类,分别是语义标注,点位标注,区域标注,序列筛选标注;其次优化了其任务上传过程,并对每种类型的标注任务进行了一对一的交互界面模板设计,使得用户能够通过这一整套交互方案完成不同类型的图像标注任务。

为了对该交互方法进行较为全面的评估和分析,本文将通过以下四个具体的项目(分别是:面部表情标注项目,身体部位标注项目,交通标志标注项目,唇语标注项目)进行实验对比,这四个项目分别所属上述提到的四种类型。我们将通过现有众包平台(Amazon Mechanical Turk)以及“众研”实验平台分别来完成这



些图像标注任务,从而进行具体的评估与分析,四个项目大致信息如表 5.1 所示:

表 5.1 项目信息表

项目编号	项目名称	所属类型	任务数量	图片数量	文件总大小
P1	面部表情标注	语义识别	100	100	4MB
P2	身体部位标注	点位标注	4949	4949	250MB
P3	交通标志标注	区域标注	100	100	10MB
P4	唇语标注	序列筛选标注	3149	26000	16.2G

由于四个项目中分别对应分类好的四种不同类型的图像标注任务,对四个项目中单个任务的描述和所需标注结果的描述如表 5.2 所示:

表 5.2 各项目的任务描述

项目编号	项目名称	单个任务描述	所需单个任务的标注结果
P1	面部表情标注	包含某个人物的一副面部表情图片,对人物表情的标注	文本(描述表情的词语)
P2	身体部位标注	包含某个人物的一张全身或者半身图片,对相应的身体部位进行点位标注	部位+相应部位的像素点坐标
P3	交通标志标注	包含一张在道路上行驶的前方路况信息图,对要求的交通信号灯或者交通指示信号进行区域标注	交通标志+区域信息(形状+中心点+长/宽)
P4	唇语标注	包含由一个短视频拆解的多张单帧图片,要对要求的图片帧所在的序列位置进行标注	序列+所在位置下标

我们招募了大约 20 众包工作者参与这些图像标注项目,然后,通过以下两种解决方案分别尝试管理和完成这些任务。

**解决方案 1(S1):** 通过线上众包模式(众包平台: AMT) 能直接完成 P1 和 P3 项目中的任务,通过线下众包工作模式间接完成 P2 和 P4 中的项目(因为目前其他线上众包平台暂不支持这两类图像标注任务)。



**解决方案 2(S2):** 通过众研(我们的实验平台), 使用我们设计的交互方法来完成这一系列图像标注任务。

实验过后,我们对相关实验数据进行了收集和统计,相关数据如表 5.3 所示。下面,我们将从标注效率,用户体验等方面来依次对这些实验数据进行评估和分析。

表 5.3 相关数据表

项目 编号	完成单个任务的平均时间(单位: 秒)		用户评分 (分数 $\in [1,5]$ , 5 分代表最满意)	
	S1	S2	S1	S2
P1	6	5	4.2	4.4
P2	30	18	2.4	4.6
P3	45	32	3.8	4.3
P4	450	360	3.0	3.8

5.2 效率评估

在第三章我们进行的对比实验中,针对通过众包平台和不通过众包平台来完成图像标注任务,我们选取了一类 序列筛选标注任务来进行测试,已得出的结论是相比传统的众包工作方式,通过众包平台来完成这类图像标注任务能够提高工作效率。

在当前完成图像标注任务的问题领域中提到,面对机器学习领域需要越来越多已标注数据作为训练集,完成图像标注的效率迫切需要提高,而完成一份图像标注任务的工作时间是衡量标注效率的主要因素,因此在这一小节中会针对完成单个图像标注任务的消耗时间来进行评估对比,根据对比的结果来评估本文所针对图像标注任务所设计和实现的交互方法是否能在标注效率上达到较好的效果。

所采取的方法是随机抽取了 4 个项目中的 10 个图像标注任务,以该任务为基准分别尝试通过方案 S1 和 S2 来完成标注工作,并统计标注耗时,相应的统计结果分别如图 5.1 至 5.4 所示:

如图 5.1 所示,项目 P1 属于图像语义标注类的项目,是一类比较基础的图像标注任务,“众研”平台和 AMT 平台均直接支持完成该类别的图像标注任务,从整体来看,随着用户完成任务数量的增多,对任务和系统使用逐渐熟悉,所需的标注时间逐渐减少,这证明了熟悉平台对该任务的标注过程,能够提高完成该

类别任务的标注效率，但是之后标注时间会逐渐稳定，在总体平均值附近波动。另外通过对比 S1 和 S2 可以看出，虽然大致趋势相同，并且部分任务完成时间比较接近，但是近一半任务，S2 的完成时间比 S1 更少，这证明了“众研”平台所设计和实现的交互方法有效的提高了完成图像语义标注类任务的标注效率。

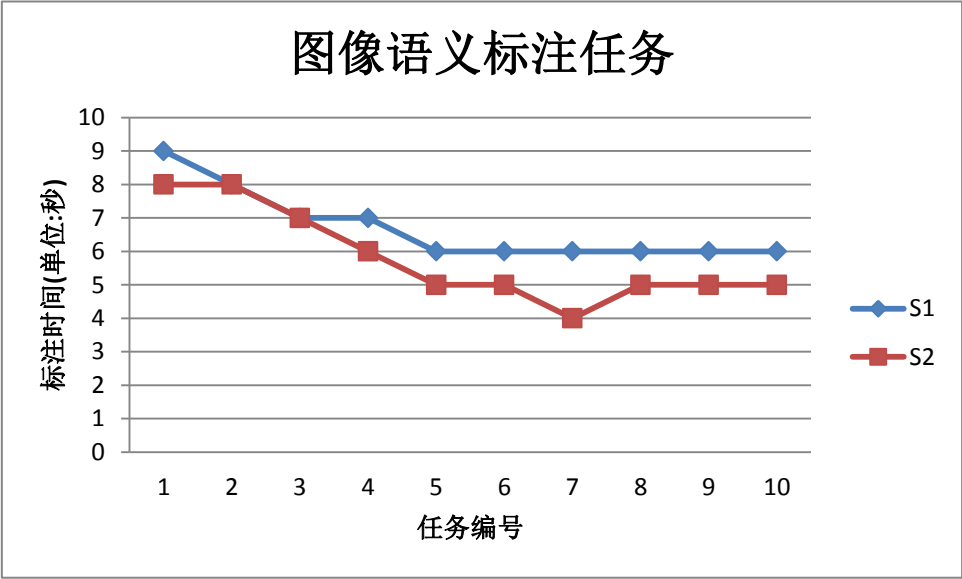


图 5.1 图像语义识别任务的平均标注时间对比折线图

如图 5.2 所示，项目 P2 属于图像点位标注类的项目，其中“众研”平台可直接完成该类别的图像标注任务，AMT 平台还未提供该类型标注任务的交互方案，但是仍然可在其平台上通过发布消息，招募众包工作者，然后通过传统的标注软件手段来完成，再通过众包平台完成结果回收和酬劳发放等过程。图 5.2 从整体上来看，方案 S2 在上述任务中的标注用时上均少于 S1 的用时。虽然在前一半的任务完成时间上，二者比较接近，但是众包工作者熟悉了标注过程之后，在后一半的任务中，S2 的完成时间比 S1 更少，这证明了“众研”平台所设计和实现的交互方法有效的提高了完成图像点位标注类任务的标注效率。

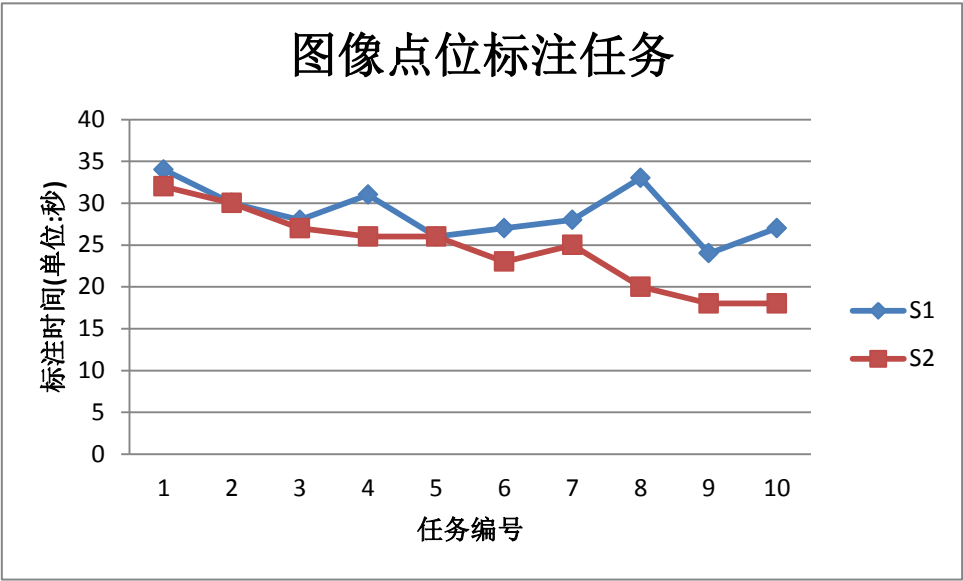


图 5.2 图像点位标注任务的平均标注时间对比折线图

如图 5.3 所示，项目 P3 属于图像区域标注类的项目，“众研”平台和 AMT 平台均可直接支持该类别的图像标注任务，从曲线总体的走势来看，方案 S1 的标注时间曲线波动较大，方案 S2 较为平稳。另外通过对比 S1 和 S2 可以看出，S2 中完成每个任务的平均时间比 S1 要少许多，这证明了“众研”平台所设计和实现的交互方法有效的提高了完成图像区域标注类任务的标注效率。

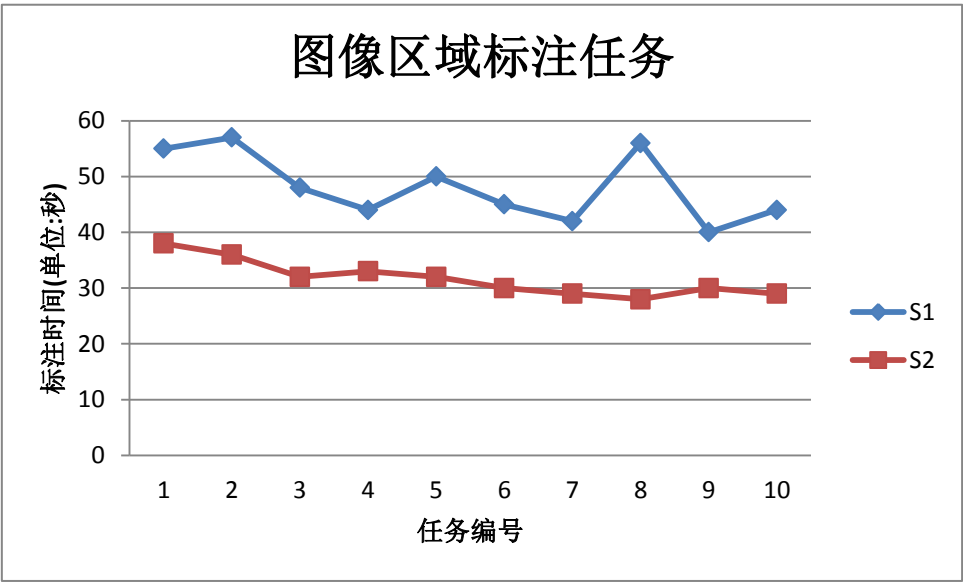


图 5.3 图像区域标注任务的平均标注时间对比折线图

如图 5.4 所示，项目 P4 属于图像序列化标注类的项目，是一种较为复杂的图像标注任务，“众研”平台可直接支持完成该类别的图像标注任务，而目前 AMT 平台还未提供该类型标注任务的交互方案，但是仍然可在其平台上通过发布消息，

招募众包工作者，然后通过传统的标注软件手段来完成，再通过众包平台完成结果回收和酬劳发放等过程。从整体上来看，S2 在上述任务中的标注平均用时上均少于 S1 的用时。这证明了“众研”平台所设计和实现的交互方法有效的提高了完成图像序列化标注类任务的标注效率。

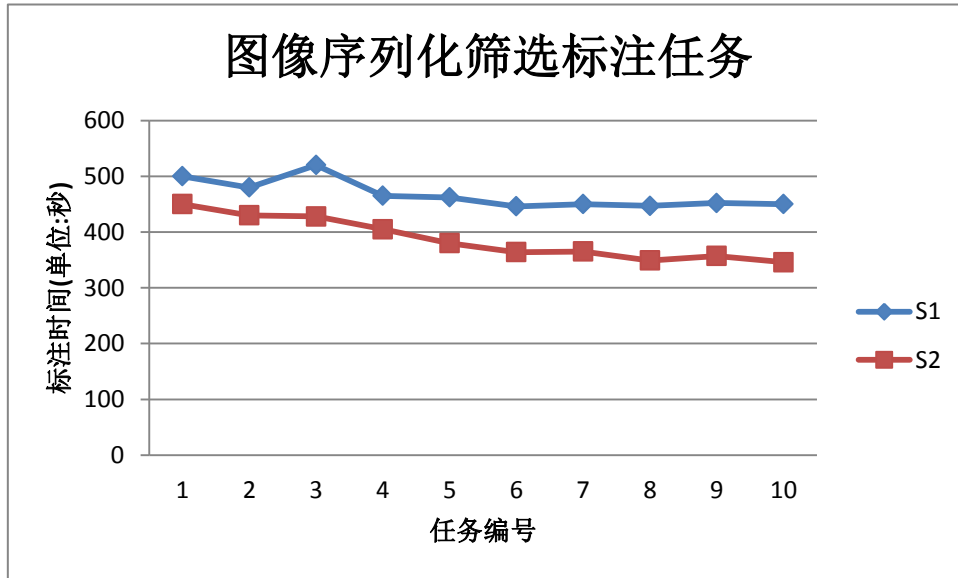


图 5.4 图像序列化标注任务的平均标注时间对比折线图

总体上，由上述四幅图以及对应的分析中可以初步得出以下结论。

在文中分类提出的四种图像标注类型的任务上，目前市场上的众包平台未完全分类并适应不同的解决方案，只能够解决部分简单的标注类型，对于复杂的标注任务类型，只能通过线上招募搭配线下标注软件方式来完成，本文基于“众研”平台所设计和实现的交互方法能够弥补这方面的不足：首先，扩充了图像点位标注和图像序列化筛选标注两种较为复杂的图像标注类型，使得四类图像标注任务都能够顺利的在“众研”平台上被解决，而且，通过对标注过程的优化，进一步减少了各类型任务的标注时间，总体上提高了四类图像标注任务的标注效率，本文基于“众研”平台而设计的交互方法，在完成图像标注任务的效率上，要略优于现有的一些线上众包平台以及传统的众包工作模式，从而更快速地制造已标注数据，提供给任务需求方。

### 5.3 用户体验评估

如问题领域中所述，由于图像任务从任务需求方传递到众包工作者的过程比较繁琐，并且当前标注过程不够简练，导致平台的两类用户缺乏一个良好的用户体验，因此在面对这四类图像标注任务的时候，用户对于当前众包平台与“众研”

平台上提供的工作方案的评分则是用户体验的一个重要指标。用户通过方案 S1 和方案 S2 完成这四类图像标注项目后，分别给出的平均评分如图 5.5 所示：

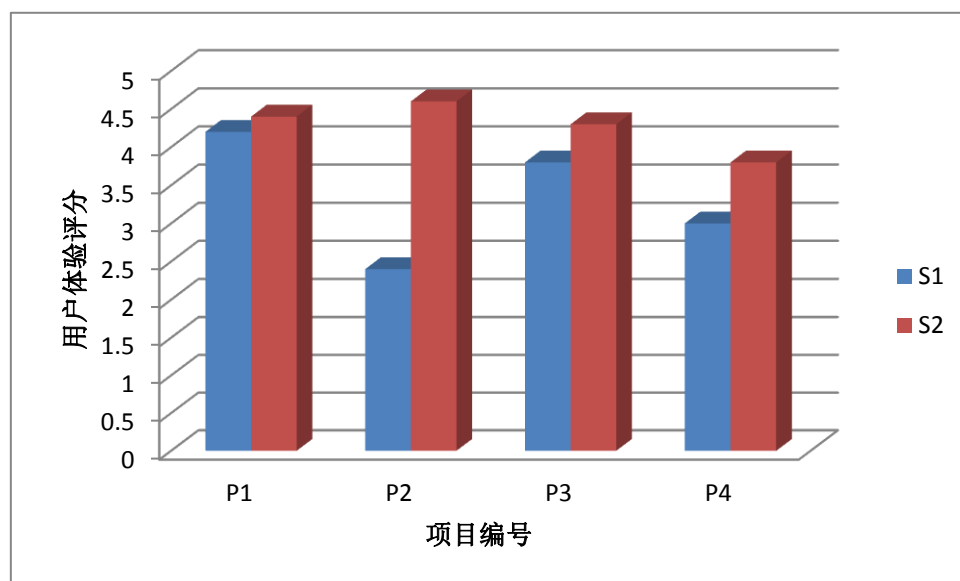


图 5.5 用户评分柱状图

在上述柱状图中，总体上来讲，完成四类图像标注项目后，方案 S2 所获得的用户平均评分均高于方案 S1，其中在项目 P2 和 P4 中(也即图像点位标注和图像序列化筛选标注任务)增幅较大，因为目前其他的线上众包平台暂时不能直接支持这两种类型的图像标注任务，仍旧只能通过线上联系和募集众包工作者+线下标注软件的形式来完成；相比之下通过方案 S2：在“众研”平台上也能直接完成这两类图像标注任务，更加简捷，从评分上来看，比较受用户欢迎。原因大致有以下几点：“众研”平台提高了工作效率，并且减少了对本地资源的占用率。马先生是我们其中的一位参与者，在进行 P4 项目时，他深有感触：

当我在本地计算机上，通过传统线下的工作方式，利用图像浏览软件工作时，因为不同人物的图像被存储在不同的文件夹，当我完成了一个任务，准备做下一个任务的时候，我需要时间来回忆我之间上一个完成的任务是哪个文件夹，然后才开始下一个任务，但是在众研平台上，它会自动显示一个未完成的任务，直到任务完成。同时呢，它能在一个页面上显示多张图片，使我可以快速的浏览图片，并在同一页的文本记录框中写下答案，而不是切换到其他窗口去记录，节约了很多的时间。

另外，项目 P1 和 P3 中的任务(也即图像语义标注和图像区域标注任务)虽然能同时其他众包平台上(如：AMT 平台)和“众研”平台上直接完成，但是，在用户体验评分上，“众研”平台要略高于 AMT 平台，这也说明了本文基于“众研”平台所设计和实现交互方法在解决图像标注任务上，用户体验要略优于当前的一些众包平台所提供的交互方法。因为在界面交互设计的过程中，我们融合了用户参

与式设计理论，深入到一些用户当中，甚至是一些特殊群体(例如残疾人)，因为一开始用户对自己所想要的交互效果并不是很清晰，通过不断的设计呈现，用户反馈，再设计，再反馈这样一个迭代的参与式设计过程，为每类图像标注任务的交互过程进行优化设计，使得任务上传和标注过程更加简单和方便。吴女士也是我们其中的一位参与者，她在项目 P3 的执行过程中，有以下体会：

*相比AMT平台，“众研”平台提供解决图像标注任务的交互方法更加简便和快捷，本身我是一名手脚不够灵活的残疾人，完成一些较复杂或者精度要求高的操作对我来讲比较费力，会更偏爱一些简单的任务和标注操作更简便的方式，这会让我更有自信去完成这些任务，另一方面，自信乐观的心态也会提高我的工作效率。*

## 5.4 本章小结

本章针对“众研”平台中针对图像标注任务所设计和实现交互方法是否能够更好的解决绪论中所提出的问题与挑战，进行了相关的实验对比分析：首先，择取了四种不同类型的图像标注项目，分别对应四类图像标注任务；然后，分别通过两种解决方案(S1 和 S2)去完成相应的标注任务；在整个过程中记录了相应任务的标注时间，定量的去分析用户的标注效率，并通过问卷和访谈的形式，定性的分析完成任务过程中的用户体验。通过曲线图以及柱状图等方式进行可视化展示，并对具体情况进行了相应的分析描述。验证了本文提出的交互方法比当前的许多众包平台上所提供的方案能更好的支持用户解决图像标注任务。

## 第六章 总结与展望

### 6.1 论文总结

本文针对目前传统的众包工作方式和现有的众包平台在面临快速提供种类丰富的图像已标注数据作为机器学习领域所需的训练数据集的过程中存在的不足,以在众包平台上完成图像标注任务为应用场景,以优化图像标注任务的交互方法,提高标注效率和用户体验为研究目标。首先通过调研机器学习领域对图像标注训练数据的需求,对图像标注任务进行了分类。通过在 pybossa 框架下搭建原型系统“众研”,基于用户参与式设计理论,以及展开的关于使用众包平台来进行众包工作的对比性研究中得出的相关结论作为设计依据,设计了一套针对解决分类得出的图像标注任务的交互方法,并在原型系统上将其实现和优化。最后通过实际的项目,随机选取了部分图像标注任务,并邀请了相应用户参与,进行了实验,通过实验数据分析得出,本文基于“众研”平台,针对图像标注任务的交互方法,首先比传统的众包工作方式能更好的胜任图像标注任务,并且相比较当前的一些众包平台能够更全面的解决图像标注任务,提高了用户的标注效率和体验。

本文的工作主要包括以下三个部分:

- (1) 通过对当前众包平台上进行的图像标注任务进行统计,然后从机器学习领域和图像识别领域中对已标注图像数据的需求进行相关调研,结合二者,对当前主要需要被解决的图像标注任务进行抽象和分类。在本文中,构建了图像标注任务的分类,分别是:图像语义标注任务,图像点位标注任务,图像区域标注任务,图像序列化筛选任务。
- (2) 搭建原型系统“众研”,展开了众包工作模式的对比性研究,即通过两种不同的众包工作模式来完成实际的图像标注任务,一种是传统的线下众包工作模式(通过即时通讯软件+相应的图像辅助软件+文本记录软件等),另一种则是目前流行的众包平台的工作模式,通过对比分析得出了利用众包平台在解决图像标注任务时的优缺点,为本文后续对交互方法的设计和原型系统的优化提供了比较重要的设计准则和原理启示。
- (3) 基于用户参与式设计理论,并结合之前对比性研究得出的相关设计准则,对分类得出的四类图像标注任务进行了交互方法和算法设计,生成一对一的交互模板,最后,通过实验数据验证本文设计和实现的交互方法能有效的提升众包工作者(尤其是残疾人等弱势群体)完成图像标注任务时

的标注效率和用户体验。

## 6.2 工作展望

本文的研究动机主要有两点：一方面是为了优化任务方和众包工作者在众包平台中完成图像标注任务工作效率和用户体验，从而更加快速和全面的为机器学习等科研领域提供数据训练集；另一方面，利用网络众包工作模式完成大数据标注工作，也能够为弱势群体(尤其是残疾人群体)提供新的就业和增加收入的机会。

本文基于“众研”平台所提出的交互方法能够有效的匹配已分类的四种图像标注任务，有效的针对用户特性提高他们在处理图像标注任务过程中的工作效率和用户体验，但是仍然存在一些不足，在未来需要在以下几方面做进一步的补充和完善：

- (1) 扩展更加多样化的标注交互方式，为不同种类的用户群体提供个性化的交互选择。因为参与众包工作的用户群体数量大，且他们之间的差异也较大，由于自身条件的不同，所以需要匹配更加个性化、差异化的交互方式，让他们选择适合自己的方式参与网络众包标注工作，从而进一步提高用户体验。
- (2) 优化图像标注任务的分类和交互模板。本文中针对当前众包平台和图像识别，机器学习等领域对于图像标注数据的需求调研，进行分类而设计的交互模板，虽然能够解决大部分对于图像标注数据的需求，但是仍然会有所遗漏。因此需要提升交互模板的扩展性，从而动态的适配更多种类的图像标注任务。



## 参考文献

- [1] Howe J. The Rise of Crowdsourcing [J]. Wired magazine, 2006, 14(6):4.
- [2] 吴俊, 崔昊哲, 赵嘉琪. 众包平台任务分类与交易方式匹配研究——基于国内外众包网站的探索[J]. 科技进步与对策, 2015(14):6-11.
- [3] Aniket Kittur, Jeffrey V. Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The Future of Crowd Work. In Proceedings of the 2013 conference on Computer supported cooperative work (CSCW '13). ACM, New York, NY, USA, 1301-1318
- [4] Von Ahn, L. and Dabbish, L. Labeling images with a computer game. Proceedings of the SIGCHI conference on Human factors in computing systems, ACM (2004), 319–326.
- [5] Benkler, Y. The wealth of networks: How social production transforms markets and freedom. Yale Univ Pr, 2006.
- [6] Malone, T.W., Yates, J., and Benjamin, R.I. Electronic markets and electronic hierarchies. Communications of the ACM 30, 6 (1987), 484–497.
- [7] Quinn, A.J. and Bederson, B.B. Human computation: a survey and taxonomy of a growing field. Proc. CHI '11, (2011).
- [8] Trist E L. The evolution of socio-technical systems: A conceptual framework and an action research program[J]. Occasional Paper No, 1981, 2.
- [9] Kulkarni A, Gutheim P, Narula P, et al. MobileWorks: Designing for Quality in a Managed Crowdsourcing Architecture[J]. IEEE Internet Computing, 2012, 16(5):28-35.
- [10] Ackerman M, Gu N, Ding X, et al. Large-scale Collaborative Projects to Affect Societal Change[C]//Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion. ACM, 2016: 544-549.
- [11] Shiri Azenkot, S. Prasain, A. Borning, E. Fortuna, R.E. Ladner, and J.O. Wobbrock. 2011. Enhancing Independence and Safety for Blind and Deaf-blind Public Transit Riders. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11), 3247-3256.
- [12] Kathryn Zyskowski, Meredith Ringel Morris, Jeffrey P. Bigham, Mary L. Gray, and Shaun Kane. 2015. Accessible Crowdwork?: Understanding the Value and Challenge of Microtask Employment for People with Disabilities. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW'15), 1682-1693.
- [13] [Online]. Available: <http://www.k68.cn>
- [14] 高云, 崔金红. 国内众包网站发展现状研究[J]. 电子商务, 2015(5):4-5.
- [15] [Online]. Available: <http://www.zhubajie.com>

- [16] [Online]. Available: <http://www.680.com>
- [17] [Online]. Available: <http://www.bangcn.com>
- [18] [Online]. Available: <http://www.taskcn.com>
- [19] [Online]. Available: <http://www.ermoo.com>
- [20] [Online]. Available: <http://www.datatang.com>
- [21] [Online]. Available: <http://www.chinacrowds.com/>
- [22] Yan, X., Ding, X., & Gu, N. Crowd work with or without crowdsourcing platforms. In Computer Supported Cooperative Work in Design (CSCWD), 2016 IEEE 20th International Conference on (pp. 56-61). IEEE. (2016).
- [23] [Online]. Available: <https://www.innocentive.com/>
- [24] 马卫, 方丽, 屠建洲. 从外包到众包的商业模式变革及启示[J]. 商业时代, 2010(1) : 23-24.
- [25] [Online]. Available: <http://www.istockphoto.com/>
- [26] [Online]. Available: <https://www.wikipedia.org/>
- [27] [Online]. Available: <https://www.mturk.com/mturk/welcome>
- [28] [Online]. Available: <https://www.upwork.com/>
- [29] [Online]. Available: <http://www.crowdflower.com/>
- [30] 谭婷婷, 蔡淑琴, 胡慕海. 众包国外研究现状[J]. 武汉理工大学学报, 2011(2): 276—279.
- [31] [Online]. Available: <http://www.stitch.com/>
- [32] [Online]. Available: <http://a.nnotate.com/image-annotation.html>
- [33] Russell B C, Torralba A, Murphy K P, et al. LabelMe: A Database and Web-Based Tool for Image Annotation[J]. International Journal of Computer Vision, 2008, 77(1):157-173.
- [34] [Online]. Available: <http://pybossa.com/>
- [35] Dewalt K M, Dewalt B R. Participant observation : a guide for fieldworkers[J]. Participant Observation A Guide for Fieldworkers, 2011.
- [36] Barbara B. Kawulich. Participant Observation as a Data Collection Method[J], 2005
- [37] Robert H. Gault. A history of the questionnaire method of research in psychology[J]. The Pedagogical Seminary, 1907, 14 (3) : 366-383
- [38] Ian. Questionnaire design : how to plan, structure and write survey material for effective market research[J]. 2015.
- [39] Irving Seidman. Interviewing as qualitative research: A guide for researchers in education and the social sciences[M]: Teachers college press, 2013
- [40] Jonathan Lazar, Jinjuan Heidi Feng, Harry Hochheiser. Research Methods in Human-Computer Interaction[J]. Handbook of Human-Computer Interaction (Second Edition), 2005: 203-227
- [41] Erickson T, Kellogg W A. Social translucence: an approach to designing systems that support social processes[J]. ACM transactions on computer-human interaction (TOCHI), 2000, 7(1): 59-83.

- 
- [42] Alpaydin E. Introduction to machine learning[M]. MIT press, 2014.
- [43] Gong Y, Leung K H T, Toshev A T, et al. Ranking approach to train deep neural nets for multilabel image annotation: U.S. Patent 9,552,549[P]. 2017-1-24.
- [44] Design and anthropology[M]. Routledge, 2016.
- [45] 李亿豪. 互联网+:创新2.0下互联网经济发展新形态[M]. 中国财富出版社, 2015



## 发表论文和科研情况说明

### 攻读硕士期间发表的学术论文

[1] Yan X, Ding X, Gu N. Crowd work with or without crowdsourcing platforms[C]// IEEE, International Conference on Computer Supported Cooperative Work in Design. IEEE, 2016:56-61.

[2] Yan X, Ding X, Gu N. An Interactive Solution for Handling Image Annotation Tasks on Crowdsourcing Platforms, 中国云计算学术大会, 2016. (Accepted)

### 攻读硕士期间参与的科研项目

[1] 上海市残疾人辅助器具资源中心项目：重症残疾人无障碍自理智能化控制与康复提醒系统, 2014.09-- 2016.10.

[2] 面向残疾人的强耦合众包协同交互机制的研究，国家自然科学基金重点项目，项目编号 61672167, 2015.06-- 2017.06



## 致 谢

转眼之间，我已经在复旦大学计算机学院生活和学习了三年，这三年无疑是我人生中非常重要的一段时光。因为它，使我的内心感觉更加充实和美好。2014年9月，我来到复旦，以一名硕士研究生新生的身份加入了协同信息与系统实验室，从一开始对学术研究和工程项目的茫然和无知，到逐渐找到方法去分析和解决遇到的问题，这一路上，虽然布满了很多的荆棘和坎坷，但同时也得到了许多贵人的援助，不管是在生活还是学习上，他们真诚的帮助与陪伴促使我顺利的度过了这段美好的旅程。毕业季马上就要来临，我谨在此，向他们表达我心中蕴含的感激之情。

首先，非常感谢我的研究生导师丁向华副教授，本论文是在丁老师悉心的指导下逐渐完成的。在这三年的学习生涯里，丁老师如良师益友般地在各个方面教会了我许多。在学习和科研方面，丁老师秉持严谨的治学态度以及精益求精的作风，促使我踏踏实实的在学术研究的道路上慢慢前行，从科研项目的实施与发布，到论文方向的选定与完成，都离不开她耐心而又实用的指导。在日常生活中，丁老师平易近人，随和开朗的性格，简单而又富含温度的话语，给予了我很多鼓励和信心。非常感谢丁老师在各个方面给予我的帮助。

其次，论文的完成也离不开顾宁教授的指导与帮助，感谢顾老师为我们提供了一个良好的学习环境，在这三年的实验室生活中，我不仅从实验室项目锻炼了自己的工程设计能力，也从中提高了自己的学术科研能力，顾老师严谨求实的作风，对实验室师生的生活和学习的关心，是我努力学习的方向。同时，也十分感谢实验室卢老师和刘老师，卢老师在论文开题，中期，例会讨论等过程中，都给予了我十分宝贵的建议和指导，刘老师在平时生活和研究生事务中给予了我很多的帮助。感谢你们对我硕士生涯的栽培与照顾。

还要感谢实验室的师兄师姐们，特别感谢孙玉灵师姐，李果师兄，刘鹏师兄，张鹏师兄，谢谢你们在我遇到科研和项目难题时为我伸出的援助之手。同时也感谢陪我一起度过三年的许家华，罗东亮，杨达一，吴娥英，程沛，我们相互鼓励，共同进步。感谢协同信息与系统实验室的全体成员，在这个学习氛围浓厚，成员彼此关心的大家族中，我变得更加乐观、更加强壮、更加富有正能量。

感谢我最亲爱的父母，是你们在背后默默的支持，无时无刻的关心问候，使我在脆弱的时候总能坚定自己前行的方向，感谢你们让我更好的茁壮成长。

最后，感谢所有参与评审我论文的各位专家与老师，谢谢你们提出的宝贵建议。



# 复旦大学

## 学位论文独创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。论文中除特别标注的内容外，不包含任何其他个人或机构已经发表或撰写过的研究成果。对本研究做出重要贡献的个人和集体，均已在论文中作了明确的声明并表示了谢意。本声明的法律结果由本人承担。

作者签名：\_\_\_\_\_ 日期：\_\_\_\_\_

# 复旦大学

## 学位论文使用授权声明

本人完全了解复旦大学有关收藏和利用博士、硕士学位论文的规定，即：学校有权收藏、使用并向国家有关部门或机构送交论文的印刷本和电子版本；允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其它复制手段保存论文。涉密学位论文在解密后遵守此规定。

作者签名：\_\_\_\_\_ 导师签名：\_\_\_\_\_ 日期：\_\_\_\_\_