西 南 交 通 大 学

本科毕业设计（论文）

机器学习边缘架构部署的研究与实现

年 级： 2016 级 \_\_

学 号： 2016113705

姓 名： 张羽晨

专 业： 软件工程

指导教师： 滕飞

二零二零年五月

西南交通大学

本科毕业设计（论文）学术诚信声明

本人郑重声明：所呈交的毕业设计（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。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者签名：

日期： 年 月 日

西南交通大学

本科毕业设计（论文）版权使用授权书

本毕业设计（论文）作者同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权西南交通大学可以将本毕业设计（论文）的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本毕业设计（论文）。

**保密**□，在 年解密后适用本授权书。

本论文属于

**不保密**☑。

（请在以上方框内打“🗸”）

作者签名： 指导教师签名：

日期： 年 月 日 日期： 年 月 日

院系 信息科学与技术学院 专 业

年级 2016级 姓 名

题目

指导教师

评 语

指导教师 (签章)

评 阅 人

评 语

评 阅 人 (签章)

成 绩

答辩委员会主任 (签章)

年 月 日

**毕业设计（论文）任务书**

班 级 软件2016-3班 学生姓名 学 号

发题日期：2019年11月22日 完成日期：2020年5月22日

题 目

1、本论文的目的、意义

2、学生应完成的任务

3、本论文与本专业的培养目标达成度如何？（如在知识结构、能力结构、素质结构等方面有哪些有效的训练。）

此为计算机科学与技术专业的最新毕业要求，其他专业应替换为各自专业的毕业要求！

本论文支撑本专业以下毕业要求的达成：**（1）**能够通过查阅和分析文献，为计算机系统及工程的问题求解寻找方案，并认识到所求解的问题具有多种可能的解决途径（指标点2.3）；**（2）**能够针对特定需求确定目标，设计计算机系统框架、组成模块，合理组织/存储数据，基于适当的模型进行系统设计与实现，并体现一定的创新意识（指标点3.3）；**（3）**能够在解决方案中从技术、非技术（如经济、社会、健康、安全、法律、文化以及环境等）角度，对设计方案的可行性进行评价和分析（指标点3.4）；**（4）**能够采用科学方法对计算机系统及工程问题进行研究，通过实验对比、文献综合、归纳整理得到合理有效结论，并对其进行规范表述（指标点4.3）；**（5）**能够利用开发环境和工具，对计算机系统及工程问题进行模拟仿真和数据分析（指标点5.3）；**（6）**能识别、分析、评价特定需求的计算机系统在设计和实现中对社会、健康、安全、法律以及文化的影响，并明确自己应承担的责任（指标点6.2）；**（7）**能够评价计算机系统设计、开发、运行和维护对环境保护和社会持续发展的影响（指标点7.2）；**（8）**能够通过口头、文稿、图表等方式、陈述和表达自己的观点，能够就计算机系统及工程问题与同行和相关人员进行交流（指标点10.1）；**（9）**能够根据对工作内容和过程的记录与整理，撰写技术报告和设计文稿、陈述发言或回应质询（指标点10.2）；**（10）**了解计算机系统工程管理原理与经济决策方法，理解计算机系统项目的组织模式和实施过程，掌握项目管理原理和内容（指标点11.1）；**（11）**正确认识自主学习的必要性和重要性，认识到本专业是一个发展迅速的学科，具有自主学习和终身学习的意识（指标点12.1）；**（12）**具备自主学习新技术和新方法的能力，能够通过学习不断提高、适应信息技术和职业的发展（指标点12.2）。

4、论文各部分内容及时间分配：（共17周）

第一部分 ( 周)

第二部分 ( 周)

第三部分 ( 周)

第四部分 ( 周)

第五部分 ( 周)

评阅及答辩 (1周)

备 注

指导教师： 2019年11月22日

审 批 人： 2019年11月22日

本页所处页数如为**偶数**，则保留；否则，如为**奇数**则删除此页。

摘 要

Web信息的爆炸性增长使Internet成为我们获取信息资源的重要途径，而在全球一体化的今天，人们对翻译质量和翻译速度的要求也日趋严格。网络上拥有大量的双语对照信息，而传统的搜索引擎无法对其进行充分的利用。因此，本文从搜索技术和翻译技术上进行研究，提出了一个基于双语翻译的搜索引擎系统。

注：

(1) 本页前面页数应为偶数，以保证此页双面打印时为正面！

(2) 至少500字。

(3) 推荐的写作段落顺序及内容如下：

①第一自然段（2~4行）：选题的提出；

②第二自然段（2~4行）：论文研究内容总概；

③第三自然段（10~12行）：描述论文所做具体内容以及其中的一些亮点（如，依据什么提出或改进了什么算法或模型，利用何方法或技术解决了什么关键问题，理论分析或仿真实验表明结果如何 等等）。

以下为示例中文摘要的**前2个自然段**：

本文利用Heritrix和Lucene工具，在计算机辅助翻译的基础上利用搜索引擎技术，实现检索具有双语对照信息的网页。本文主要研究如何从海量信息库中爬取网页资源，设计出识别双语网页和提取双语语料的方法，并构造出合适的索引器和检索器，通过用户接口将网页信息输出给用户。

……

关键词：词1；词2；词3；词4；词5

注：① 关键词与摘要正文之间空一行；

② 关键词应按其涵义外延从大到小顺序排，列3-5个。

Abstract

The explosive growth of Web information makes Internet be an important way to obtain information resources for us. Because of today’s global integration, the quality and speed of translation work are becoming stricter and stricter. There is a large amount of bilingual information in the network, but traditional search engines cannot make full use of them. Therefore, this thesis studies search and translation technologies and proposes a search engine system based on bilingual translation.

注：特别注意此部分不能用中文标点符号，而应用英文标点+空格（一个）来分隔。

以下为示例英文摘要的**前2个自然段**：

This thesis realized searching Web pages that have bilingual information, and combined computer-aided translation technology with search engine technology with Heritrix and Lucene tools. This thesis mainly researches how to crawl Web resources from massive information, and designs a method for recognizing bilingual Web pages and extracting bilingual corpora. The search engine implements in this thesis will output Web information to users through the user interface with the help of indexers and searchers.

……

**Keywords:** kw1; kw2; kw3; kw4; kw5

注：① Keywords与英文摘要正文之间空一行；

② 与中文关键词对应排列。

目 录

[第1章 绪 论 1](#_Toc452574320)

[1.1背景与意义 1](#_Toc452574321)

[1.2国内外发展（应用）现状 1](#_Toc452574322)

[1.3论文所做工作及思路 2](#_Toc452574323)

[1.4论文章节安排 2](#_Toc452574324)

[第2章 标题 4](#_Toc452574325)

[2.1一级节标题 4](#_Toc452574326)

[2.1.1二级节标题 4](#_Toc452574327)

[2.1.2二级节标题 5](#_Toc452574328)

[2.1.3二级节标题 5](#_Toc452574329)

[2.2一级节标题 5](#_Toc452574330)

[2.2.1二级节标题 5](#_Toc452574331)

[2.2.2二级节标题 5](#_Toc452574332)

[2.3本章小结 5](#_Toc452574333)

[第3章 标题 6](#_Toc452574334)

[3.1一级节标题 6](#_Toc452574335)

[3.1.1二级节标题 6](#_Toc452574336)

[3.1.2二级节标题 6](#_Toc452574337)

[3.2一级节标题 6](#_Toc452574338)

[3.2.1二级节标题 6](#_Toc452574339)

[3.2.2二级节标题 6](#_Toc452574340)

[3.3本章小结 6](#_Toc452574341)

[第4章 标题 7](#_Toc452574342)

[4.1一级节标题 7](#_Toc452574343)

[4.1.1二级节标题 7](#_Toc452574344)

[4.1.2二级节标题 7](#_Toc452574345)

[4.2一级节标题 7](#_Toc452574346)

[4.2.1二级节标题 7](#_Toc452574347)

[4.2.2二级节标题 7](#_Toc452574348)

[4.3本章小结 7](#_Toc452574349)

[第5章 标题 8](#_Toc452574350)

[5.1一级节标题 8](#_Toc452574351)

[5.1.1二级节标题 8](#_Toc452574352)

[5.1.2二级节标题 8](#_Toc452574353)

[5.2一级节标题 8](#_Toc452574354)

[5.2.1二级节标题 8](#_Toc452574355)

[5.2.2二级节标题 8](#_Toc452574356)

[5.3本章小结 8](#_Toc452574357)

[结 论 9](#_Toc452574358)

[致 谢 10](#_Toc452574359)

[参考文献 11](#_Toc452574360)

[附录1 标题 13](#_Toc452574361)

[附录2 标题 14](#_Toc452574362)

第1章 绪 论

注：(1) **本页前面页数应为偶数**，以保证此页双面打印时为正面！

(2) 正文中文用宋体，英文与数字用Times New Roman，均为小四号。

(3) 论文篇幅要求从第1章至参考文献原则上40页。

1.1背景与意义

注：所有章节标题后不能带标点符号。

1.2国内外发展（应用）现状

注：本节应标注文献引用！尽量**避免**直接引用方式，且尽量**用自己理解后的话**来叙述！引用的**上标**用**小四**号字，**Times New Roman**字体[1]，多篇文献的标注如[2, 5-8]。当提及的参考文献为文中直接说明时，其序号应该用小四号字与正文排齐，如“由文献[8, 10-14]可知”。文中其它章节如有引用亦应标注！注意：全文**不能在标题上标注**！

1.3论文所做工作及思路

注：本节应从以下方面叙述：本论文将要做什么内容，准备分几步来做，各步分别准备怎么做，其中可能会涉及什么关键问题或技术，准备如何解决等等。

1.4论文章节安排

论文共分xx章。

注：以下分章简述各章的内容。

第1章……

第2章……

……

注：

(1) 文中的图、表不要被分页断开！表如被分到两页，则应对分在另一页的表的右上方写上“续表”。图与图名不能分到两页上。图应尽量紧凑些。表的左右两个边框应隐去。图表应分别编号（不能混编），且均以“图 章号-顺序号、表 章号-顺序号”形式按章编号（例如，图1-1，表示第1章的第1张图；图2-1表示第2章的第1张图；表1-1表示第1章的第1张表；表2-1表示第2章的第1张表）。引用别人的图或表（亦应在文中标注引用），尽量自己重新绘制，不要直接复制。图表中文字（包括图名、表名）字号建议用5号字。

以下的表1-1和图1-1为表和图的示例。

表1-1 Corpus表

|  |  |  |
| --- | --- | --- |
| **列名** | **数据类型** | **说明** |
| Cid | Int | 自增型单页语料编号，初值为1 |
| Url | Varchar(200) | 网页地址 |
| Ccn | Text | 单页互译文本的中文句子 |
| Cen | Text | 单页互译文本中的英文句子 |



(2) 按照GB3100～3102及GB7159-87的规定，物理量符号、物理常量、变量符号（如：*a*(*t*) , (*i*-1)*Th**t*<*iTh ,m ,n*）用斜体，计量单位（如：km、KB）等符号均用正体。

(3) 公式居中写，大小与正文相同，公式末不加标点，序号按章节编排，如有“假定，解”字样，文字空两格写，若有对公式变量的说明，以分号结束。公式中用斜线表示“除”的关系时应采用括号，以免含糊不清，如1/(*b*cos*x*)。通常“乘”的关系在前，如*a*cos*x*/*b*而不写成(*a*/*b*)cos*x*。具体格式如下：

 (1-1)

式中，*R*表示幅度；*θ*表示相位；*fc*表示载波频率。

(4) 正文中不能出现大段的空白部分（每章结束处除外）！

第2章 需求分析和概要设计

2.1系统设计目标

本系统旨在实现对基于计算机视觉技术的智能应用的边缘架构部署。通过协调数据预处理、数据上传、模型训练、模型下放、实时推断的过程，使该智能应用，能够在云端高性能服务器进行机器学习模型的训练过程，而在边缘端进行实时推断和数据收集。在方便智能应用的维护、更新的同时，降低网络带宽占用，提高应用实时性。

基于计算机视觉技术的智能应用具有以下特征：

1. 源数据为图片、视频等多媒体，在数据量多的情况下占用较多网络带宽。
2. 模型训练过程对计算、存储性能要求高，普通服务器往往无法满足。
3. 模型运行过程则不需要过多计算资源，可以在离用户进的边缘端进行。

因此，本系统的服务程序将分为云端、边缘端，共同完成以下核心目标：

1. 能够在边缘结点进行应用数据的接收和标注数据的收集
2. 能够将边缘结点上的数据进行预处理后传送到云端服务器
3. 能够管理在云端服务器进行机器学习模型训练的过程
4. 能够协调云端服务器完成训练后向边缘服务器下发模型的过程
5. 能够在边缘服务器加载模型提供智能应用服务，实现实时推断

2.1.1

2.1.1.1三级节标题

注：节最多列到三级，但非必须，需要时才列；如内容过多，还可按条款（见下，最多3级）进行组织。如内容不太多，也可以不设三级节，直接用条款来组织。

1. 第1级条款标题（退2格，不能太长，不能以标点结束，也不能加引用标注）

（1）第2级条款内容（首行退2格，后面的文字不要悬挂缩进直接顶格换行，以标点结束）

① 第3级条款内容（首行退2格，后面的文字不要悬挂缩进直接顶格换行，以标点结束）

2.2需求分析

2.2.1用户分析

本系统的主要用户为智能应用的开发者，通过将机器学习模型接入该系统，用户可以在边缘端完成标注数据的收集并向智能应用的普通使用者提供低延迟的实时推断服务；并通过访问边缘端web页面来管理数据的预处理和上传、模型在云端的增量训练、模型下放和更新过程。

2.2.2系统功能需求

1. 边缘端系统功能需求

（1）向智能应用开发者提供网络访问接口，使开发者能够操作系统相应功能并向应用使用者提供对智能应用的服务访问。

（2）接收应用数据并进行预处理。

（3）接收标注数据处理后写入当前收集数据集文件。

（4）通过与云端通信，对边缘端本地和以及云端的数据集文件进行管理，实现以下具体操作：

① 基本文件管理：数据集文件的新建、查询、删除以及合并；

② 数据集上传与同步：获取数据集存储状态，能够由边缘端向云端同步数据集文件以及由用户本机向边缘端上传新数据集文件；

③ 设置当前收集数据集，使接收到的标注数据能够追加到指定数据集内；

（5）通过与云端通信，完成对模型训练过程以及模型文件的管理，实现以下具体操作：

① 向云端程序发出在指定已上传数据集上以指定参数开始训练的请求；

② 获取云端训练状态，向云端程序发出终止进行中的训练任务的请求；

③ 查询本地以及云端所有的可用模型详细信息，包括名称、大小、生成时间、训练及测试详情；

④ 向云端程序发出下载指定模型的请求。

⑤ 将指定模型设置为当前模型，并响应普通用户对该智能应用服务的访问，加载指定模型进行实时推断，返回推断结果。

2.2.3系统性能需求

1. 网络传输需求

本系统的实现需要依赖边缘端程序和云端程序之间的内部通信过程，因此需要设计并实现合理的自定义应用层通信协议。由于TCP是面向字节的流协议，发送的数据会经过合并优化进行发送和接收；因此在发送端和接收端分别需要设置合理的组包、解包模块，以避免“粘包问题”，即由于接收端不知道该从连接在一起的数据流中读取多长的数据而造成的接收数据解析错误。

2. 并发性需求

本系统为多端多用户的网络程序，对并发性有一定要求。在同时需要对多端进行响应时需要能够正确、异步地进行实时响应。

2.3数据对象分析

2.4系统结构和模块关系设计

边缘端：交互层、业务层、机器学习项目

云端：业务层、机器学习项目

2.5项目技术选型

边缘端：交互层、业务层、机器学习项目

云端：业务层、机器学习项目

2.6本章小结

第3章 系统详细设计与实现

3.1系统核心服务模块设计与实现

编写自定义类NetDeviceBase、FileManagerBase、CloudServer、EdgeServer类实现核心业务层逻辑。其中NetDeviceBase、FileManagerBase提供系统通用的网络通讯及文件管理服务；而CloudServer、EdgeServer在继承两个基类获取网络通讯和文件管理功能的基础上，重写连接处理函数提供对不同类型消息的个性化响应；并分别对应云端、边缘端的需求，提供直接调用的服务接口。

3.1.1网络通讯子模块

1. 应用层协议设计与实现

本文采用TLV(Type- Length -Value)结构的思想设计应用层协议。即以类型域、数据长度域和数据域的结构组成数据包，如图[]所示。其中类型域、长度域的长度往往事先约定，数据域的长度为长度域的取值。在解析到达的流数据时，每次先按固定长度读出消息类型、长度值，即可知道接下来需要读取的数据域的长度。读取后，根据类型的取值不同对不同的消息加以特定处理。

|  |  |  |
| --- | --- | --- |
| Type(n) | Length(m) | Value(Length) |

TLV结构

本文编写TLV.py模块提供TLV、TLVParser两个类，在发送端通过TLV类实现组包，在接收端通过TLVParser类实现流数据拆包解析。下面介绍其主要数据成员与方法。

1. <class ‘TLV’>

TLV类的主要任务是进行组包，其数据成员概况如表[]所示。

表1-1 Corpus表

|  |  |  |  |
| --- | --- | --- | --- |
| **命名** | **数据类型** | **初始化默认值** | **说明** |
| t\_len | <class 'int'> | 1 | 类型域所占字节数。 |
| l\_len | <class 'int'> | 1 | 长度域所占字节数。 |
| buffer | <class ' bytearray '> | bytearray() | 数据组包缓冲区。 |

其主要方法成员如下所述:

add\_obj((self, type, value)：首先调用pickles模块将python对象转换成字节数据、计算出长度，再将指定的类型、长度和字节数据转换成以TLV格式组织的字节数组追加到缓冲区。

pop\_buf(self)： 返回TLV实例当前缓冲区中的组好包的流数据，同时清空缓冲区。

TLV组包算法具体流程如图[]所示



add\_obj算法流程图



1. <class ‘TLVParser’>

TLVParser类的主要任务是进行拆包。本文提供的TLVParser类可与一个socket绑定，实现在接收数据不完整、无法解包的情况下，捕捉异常并再次接收数据。其数据成员概况如表[]所示。

表1-1 Corpus表

|  |  |  |  |
| --- | --- | --- | --- |
| **列名** | **数据类型** | **初始化默认值** | **说明** |
| t\_len | <class 'int'> | 1 | 类型域所占字节数。 |
| l\_len | <class 'int'> | 1 | 长度域所占字节数。 |
| step | <class 'int'> | 1024 | 在配置了数据源套接字的情况下，进行二次数据接收时接受的数据量上限。 |
| buffer | <class ' bytearray '> | bytearray() | 数据组包缓冲区。 |
| offset | <class 'int'> | 0 | 解析进度下标，保存对缓冲区内字节数组当前解析完毕字节后的下一个字节。 |
| socket | <class 'socket.socket'> | None | 待解析的流数据源套接字，可留空。若指定了源套接字，在缓冲区内数据被截断、因数据不完整无法解析触发TLVError时，尝试从套接字中进行再次接收，组成完整流数据再次尝试解析。 |

其主要方法成员如下所述:

\_get\_tlv (self)：按照约定的类型域、长度域长度读取类型、长度取值，依此读取数据域字节数组，再通过pickles模块恢复数据为python对象，以字典形式返回解析出的一组t，l，v值。若产生截断数据导致的解析错误则抛出TLVError。

parse\_obj (self)：调用方法获取并以字典形式返回解析出的一组t，l，v值。当捕捉到TLVError时，再次接收套接字中的数据，再次尝试解析。

TLV拆包算法具体流程如图[]所示



拆包算法流程图

2. 网络通讯工具包设计与实现

本文编写net\_device\_base包实现对系统所需常用网络功能的封装，提供模板类NetDeviceBase类和数个工具方法。以NetDeviceBase作为EdgeServer、CloudServer的父类，就可以使边缘端、云端服务类直接调用针对本系统编写的内部通讯的公共代码。

表1-1 Corpus表

|  |  |  |
| --- | --- | --- |
| **静态方法** | | |
| **方法名** | **输入参数** | **说明** |
| get\_new\_tcp\_connection  (addr) | addr - 待连接的地址。 | 新建tcp套接字，尝试连接指定地址，若连接成功则返回套接字。 |
| send\_msg(conn, type, body) | conn - 已连接的套接字。  type - 待发送的消息类型编码。  body - 待发送的消息体，可为任何python对象。 | 以指定套接字向所连接地址发送指定消息。 |
| expect\_msg(conn, type) | conn - 已连接的套接字。  type - 预期接收的消息类型。 | 等待从指定套接字接收预期消息，如超时或收到错误类型消息则报错。 |
| send\_data(conn, filename, step = 1024) | conn - 已连接的套接字。  filename - 待发送的文件路径。  step - 发送步长，默认为1024。 | 向已连接的套接字发送指定路径下全部文件数据。 |
| recv\_data(conn, filename, filesize, step=1024) | conn - 已连接的套接字。  filename - 待保存的文件路径。  filesize - 待接收的数据大小。  step - 接收步长，默认为1024。 | 从已连接的套接字接收指定大小的数据，并保存到所配置的路径下。 |
| upload\_file(conn,filepath) | conn - 已连接的套接字。  filename - 待发送的文件路径。 | 通过文件发送请求、文件接收接收确认消息沟通待发送文件大小、名称等信息后，向指定套接字发送指定文件，确保文件的成功传输。 |
| **动态方法** | | |
| **方法名** | **输入参数** | **说明** |
| tcp\_listen(self, conn\_num) | self -自身引用，用于获取配置的服务地址等。  conn\_num - 最大同时连接数。 | 新建tcp套接字并绑定自己的服务地址，开启线程不间断尝试获取新的连接；一旦获取新的连接则开启deal\_conn线程分别处理新的连接。 |
| deal\_conn(self) | self - 自身引用，用于获取配置的服务地址等。 | 由子类继承重写，不断尝试接收新的数据，当数据不为空时，调用TLV模块解包，获取消息对象，根据消息类型不同进行不同的响应。 |

3.1.2文件管理子模块

由于机器学习项目本身以及其所依赖的python库的特点，训练数据、模型参数数据数据往往直接进行文件存取。pandas和numpy模块能够方便地读取文件数据并转换成训练所需的格式。为了能够和大部分机器学习项目更好地接洽，本系统提供file\_manager\_base模块，直接对数据集、模型文件进行管理。file\_manager\_base模块包含一个FileManagerBase类，提供以下静态方法，如表所示。

表1-1 Corpus表

|  |  |  |
| --- | --- | --- |
| **静态方法** | | |
| **方法名** | **输入参数** | **说明** |
| listdir(dirpath, rdb=None) | dirpath - 磁盘文件夹路径。  rdb - 要求后缀名信息，默认为None，表示不限制。 | 返回文件夹下指定后缀名的文件的详细信息列表。 |
| get\_remote\_file\_list(addr,type) | addr - 远端服务地址。  type - 用于确定远端文件夹路径和文件后缀名要求的类型参数，如’data’，’model’。 | 获取并返回远端指定文件夹下文件的详细信息列表。 |
| merge\_list\_with\_mark(l\_e, l\_c, mark=['on\_edge', 'on\_cloud', 'on\_both']) | l\_e - 边缘端文件信息列表  l\_c - 云端文件信息列表  mark - 合并状态标记列表，默认分别为['on\_edge', 'on\_cloud', 'on\_both'] | 合并边缘端、云端对相同路径下的文件列表，并增加列表中的信息域标记它们的状态；返回合并后的信息列表。 |

3.1.3云端服务子模块

1.消息映射

云端服务类CloudServer重写父类的deal\_conn方法，将解析出的不同类型消息交由指定处理方法进行处理。映射关系如下表所示

表1-1 Corpus表

|  |  |  |
| --- | --- | --- |
| 消息类型 | 对应处理方法 | 说明 |
| SEND\_DATA\_REQ | deal\_send\_data\_req | 发送数据请求 |
| GET\_FILE\_LIST\_REQ |  | 获取文件列表 |
| GET\_FILE\_REQ |  | 获取指定文件 |
| TRAIN\_MODEL\_REQ |  | 训练指定模型 |

表中所述消息处理方法的设计阐述如下：

3.1.4边缘端服务子模块

1.消息映射

云端服务类CloudServer重写父类的deal\_conn方法，将解析出的不同类型消息交由指定处理方法进行处理。映射关系如下表所示

表中所述消息处理方法的设计阐述如下：

表1-1 Corpus表

|  |  |  |
| --- | --- | --- |
| 消息类型 | 对应处理方法 | 说明 |
| SEND\_DATA\_REQ | deal\_send\_data\_req | 发送数据请求 |
| GET\_FILE\_LIST\_REQ |  | 获取文件列表 |
| GET\_FILE\_REQ |  | 获取指定文件 |
| TRAIN\_MODEL\_REQ |  | 训练指定模型 |

2.服务接口

除去处理来自系统内部云端的消息的消息处理方法外，由于本系统是在边缘端提供管理服务的，EdgeServer类还提供一系列公共方法作为核心服务模块的对外接口，以供系统管理模块调用。

表1-1 Corpus表

|  |  |  |
| --- | --- | --- |
| **方法名** | **输入参数** | **说明** |
|  |  |  |
|  |  |  |
|  |  |  |

3.2系统管理模块设计与实现

边缘端程序由于需要为智能应用普通用户提供实施推断服务以及需要接收各种来自智能应用开发者的管理命令，需要对外提供用户友好的交互界面；本文基于web服务实现系统管理模块，提供轻量、用户友好的操作界面。

系统管理模块依赖于核心服务模块，主要任务是获取并解析用户的输入，调用核心服务模块中边缘端服务子模块相应接口，返回标准数据。

系统管理模块分为智能应用服务、数据集管理、模型管理和设置四方面的功能。以下将分别讲解四个子模块交互层、业务层的设计与实现。

3.2.1智能应用服务模块

1. 交互层

智能应用服务模块的交互层部分，旨在为本程序在边缘端上的实时推断、标注数据收集过程提供演示和测试途径；在实际应用过程中，用户可以灵活地在自行编写的程序或网页中调用该服务模块的业务层网络接口，实现高响应的实时推断服务并允许随应用过程进行标注数据收集。

在该演示模块中，系统会要求用户上传用于推断的图片，获取图片后将其显示在网页上，并调用业务层将图片进行标准化预处理和转存。随后，用户点击预测按钮，交互层调用业务层，传入处理后的图片在边缘结点上的路径并加载当前模型运行应用，预测结果将返回显示在网页上。同时，给出数据标注表单，允许用户反馈预测结果是否正确以及正确的结果。通过调用业务层“”函数，该标注将被追加到当前收集数据集文件中。

应用服务页面展示

实时推断及数据标注操作流程

2. 业务层

（1）实时推断的实现

实时推断通过调用机器学习模型的应用接口实现。【】

（2）数据标注的实现

数据标注的业务实现首先是依据训练所需数据集格式对图片数据进行标准化。以本文为例，需要将图片首先转换成48×48的灰度值矩阵，再按行序排列转换成长度为2304的数组，格式化为字符串的形式，作为‘Pixels’列的数据；而标注的文本信息转换为分类编号，格式化为字符串后作为‘Label’列数据；随后，调用pandas包将二者组合成一个DataFrame后追加至当前所设定的csv文件，完成标注数据的收集。具体流程如【】所示。

数据标注流程图

3.2.2数据集管理模块

数据集查询：查询数据集列表获取详细信息并判断数据集位置，首先会检索边缘结点本地所配置的数据集路径下的所有数据集文件，获取它们的详细信息，包括最后文件大小、数据条目、创建时间和最后修改时间；然后边缘端会进一步向云端请求云端的数据集文件和其详细信息；获取到两个数据集列表后，对其进行合并，并标注数据集文件的位置状态信息。随后将结果返回前端网页，生成表格。在最后的操作列中，加入“上传云端”和“训练”按钮，而只有对于位于云端的数据集，“训练”按钮才是可用的。

数据集上传：

在数据集上训练：

点击训练按钮，将会弹出配置窗口，用户可以输入目标训练轮次、每轮喂入数据量以及训练过程中备份模型间隔轮数。

3.2.3模型管理模块

3.2.4设置模块

3.3对外接口设计与实现示例

3.3.1

3.3.2

3.3.3

3.3.4

第4章 系统测试与发布

4.1系统测试环境

4.1.1二级节标题

4.1.2二级节标题

4.2一级节标题

4.2.1二级节标题

4.2.2二级节标题

4.3本章小结

结 论

注：应客观地总结性说明本论文已经做了哪些方面的工作，各方面又是采用什么方法/手段/技术做了哪些主要内容，取得什么结论/效果；

对之后的展望，应说明在本论文研究工作基础上，今后可进一步研究或完善的问题，或本论文所做工作还需进一步完善的地方，列2~3条即可。

总结与展望：

可以考虑探索部署单独的用户服务程序，

将用户服务程序部署在边缘端。

这样可以直接向所有提供统一的服务接口，

减少用户部署的工作量和环境带来的不确定性。

致 谢

注：

（1）查重会覆盖整篇论文（包括致谢），因此所有部分都不得抄袭、复制！

（2）除对导师感谢外，对于校外毕业设计，也需要感谢校外导师及其所在单位！

参考文献

要求：(1) 篇数≧20。

(2) 格式要规范（见下，其中蓝色为示例）。

(3) 英文文献中人名以“姓+名的缩写”方式（参见下述示例）。

(4) 文献应在文中（特别是在1.2节）引用！

(5) 序号应按文中标注出现的先后顺序排。

(6) 遵照以下各种类型文献格式标准列出参考文献，多个作者间用逗号（, ）分隔。注意各文献各项信息之间的分隔符均应使用英文半角，并加一空格。

1. 作者名. 期刊论文题目[J]. 期刊名称, 年, 卷(期): 起止页码.
2. 刘东飞, 陈智贤, 邓俊涛, 余俊. 一种辅助翻译搜索引擎的实现[J]. 计算机工程, 2007, 33(1): 216-218.
3. Zhang S., Liu L., Diao L.H. Problems on Large-scale Speech Corpus and the Applications in TTS [J]. Chinese Journal of Computers, 2010, (4): 687-696.
4. 作者名. 书名[M]. 第n版. 出版社名称, 出版年份. （注: 如为第1版则省略不写版次信息）
5. 陶宏才. 数据库原理及设计[M]. 第3版. 清华大学出版社, 2014.
6. Stallings W. Data and Computer Communications [M]. 10th ed. Pearson Education, Inc., 2014.
7. 作者名. 书名[M]. 译者. 第n版. 出版社名称, 出版年份. （注：此为译著格式）
8. Abraham S., Henry F.K., Sudarshan S. Database System Concepts. 杨冬青 等译. 6th ed. McGraw-Hill Companies, 2012.
9. 作者名. 学位论文题目[D]. 博/硕士论文. 学校, 年份.
10. 杜亚军. 搜索引擎智能行为的研究及实现[D]. 博士论文. 西南交通大学, 2006.
11. 曹桂锋. 搜索引擎中网页分类和网页净化的研究与实现[D]. 硕士论文. 武汉理工大学, 2013.
12. 作者名. 会议论文题目[C]. 学术会议论文集名称, 出版年份: 起止页码.
13. Zhang X., Dong Y. Designing and Implementing of the Webpage Information Extracting Model Based on Tags[C]. International Conference on Intelligence Science and Information Engineering, 2011: 106-111.
14. 作者名. 标准名称[S]. 标准机构名称, 年份.
15. Beser B., Droms R., Stump G., et al. The User Class Option for DHCP [S]. RFC3004. 2000.
16. 作者名. 报纸文章名[N]. 报纸名称, 年-月-日.
17. 王菲. 移动社交APP大比拼[N]. 上海金融报, 2014-12-16.
18. 作者名. 技术报告题目[R]. 机构名称, 年份.
19. EMC Corporation. Symmetric Remote Data Facility Product Description Guide[R]. EMC Corporation, 2000.
20. 作者名. 网上文章题目[EB/OL]. 网址, 查阅的年-月-日.
21. 中国报告大厅. 2015年全球智能手机用户数量预测数据分析[EB/OL]. http://www.chinabgao.com/stat/stats/ 39758.html, 2015-3-24.

文献类型和标志代码对照表（参考后须删除此表）

|  |  |
| --- | --- |
| **文献类型** | **标志代码** |
| 普通图书 | M |
| 会议录 | C |
| 汇编 | G |
| 报纸 | N |
| 期刊 | J |
| 学位论文 | D |
| 报告 | R |
| 标准 | S |
| 专利 | P |
| 数据库 | DB |
| 计算机程序 | CP |
| 电子公告 | EB |

附录1 标题

注：对需要收录于毕业设计中且又不适合在正文中书写的附加数据、资料、详细公式推导等有特色的内容，可作为附录排写。

附录2 标题