iMoments

软件架构文档

版本 1.0

修订历史记录

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **版本** | **说明** | **作者** |
| 2021/4/12 | 1.0 | 完成架构设计 | 沈玮杭 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

目录

1. 简介 4

1.1 目的 4

1.2 参考资料 4

2. 用例视图 4

3. 逻辑视图 5

3.1 概述 5

4. 进程视图 6

5. 部署视图 6

6. 实现视图 7

7. 技术视图 7

8. 核心算法设计 7

9. 质量属性的设计 7

软件架构文档

# 简介

## 目的

本文档将从构架方面对系统进行综合概述，其中会使用多种不同的构架视图来描述系统的各个方面。它用于记录并表述已对系统的构架方面作出的重要决策。

## 参考资料

无

# 用例视图

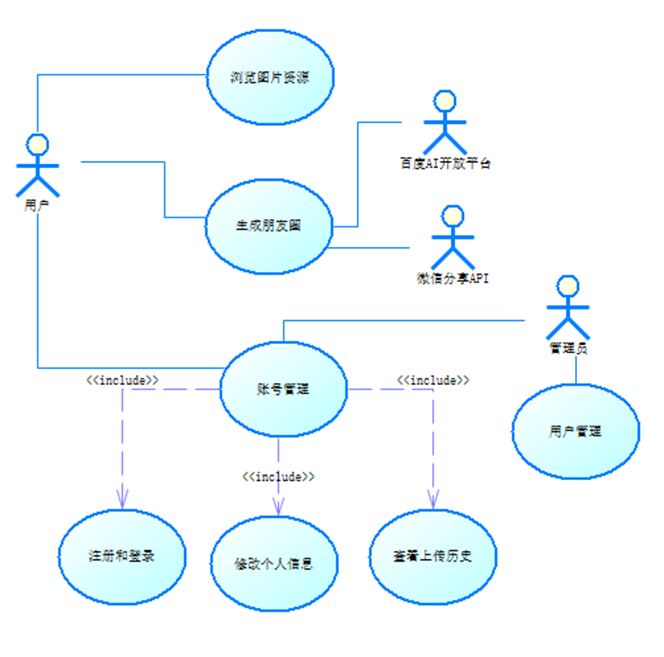


图1 用例视图

用户可以选择上传想发的图片或一段文字，iMoments可以向云端发送请求，使用百度AI开放平台提供的API提取出图片的关键要素，分析文字的主题和情感倾向。根据分析图片和文字得到的信息，云端会通过神经网络模型生成优美的诗歌散文等配文，返回给手机端呈现。

在手机端，iMoments可以根据用户输入的提示文字生成其他类型的配文，如藏头诗等。最终，手机端会将云端和本地生成的配文全部反馈给用户供用户自由选择。用户还可以浏览并保存精美的朋友圈配图。

用户还可以对生成的配文进行编辑或选择将生成的配文复制到剪切板。用户也可以选择收藏配文、图片等处理结果，并在收藏夹中查看。

# 逻辑视图

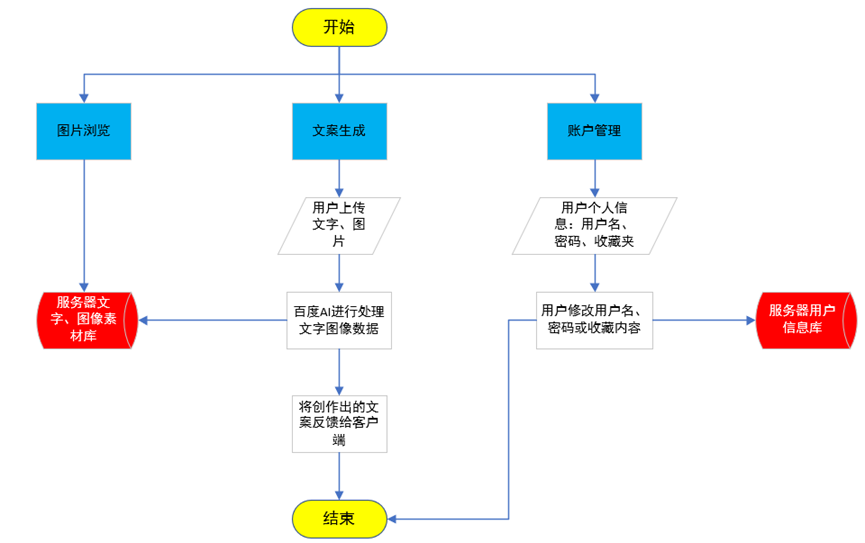


图2 逻辑视图

## 概述

iMoments的功能实现可以分解为以下几个子问题及实现途径：

① 在用户上传的图片中提取关键要素：百度云端通用物体和场景识别API。

② 提取用户输入文字中的关键主题和情感倾向：百度云端语义和情感倾向分析API。

③ 生成配文：本地 + 云端神经网络模型迭代求解。

# 进程视图

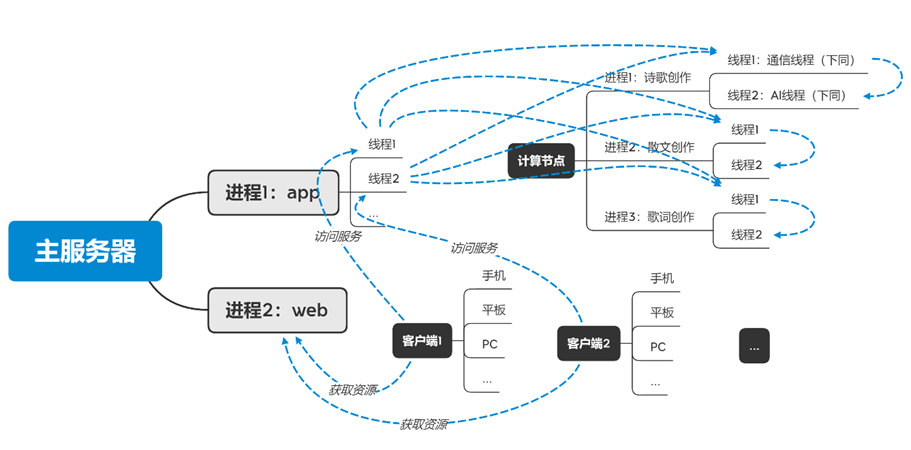


图3 进程视图

主服务器中运行了两个进程。进程1是向提供app配文创作服务的。进程2是向客户端提供Gallery图片服务的。每个客户端请求一次生成配文服务，进程1都会开出一个新的线程来处理此用户的请求。该新的线程会调度各计算节点，向合适的计算节点的三个配文创作进程的通信线程分配计算任务。通信线程向AI计算线程发送请求。计算线程将计算结果逐层返回给客户端。

# 部署视图

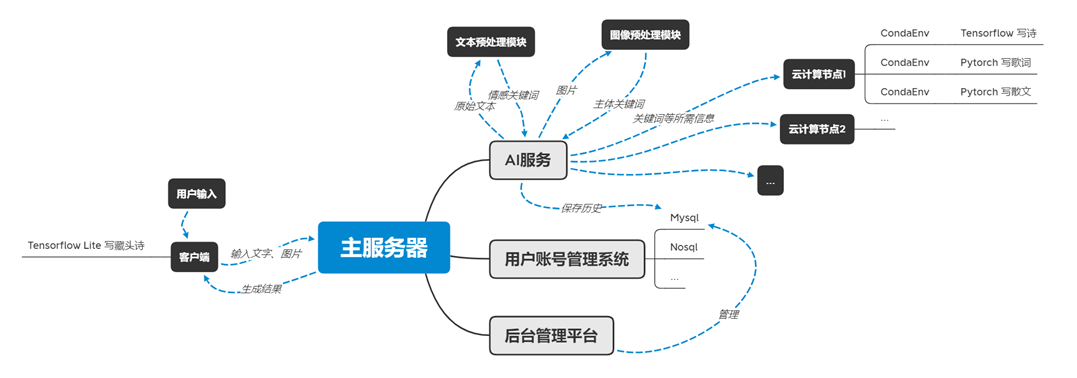


图4 部署视图

系统主要分为移动客户端和云端。其中客户端接受用户输入，在本地运行TensorflowLite模型生成藏头诗，分担云端运算压力，并把输入的文字和图片上传主服务器。主服务的AI服务模块将调用百度AI平台预处理文本和图片，提取情感倾向和图片主体关键词，并将预处理结果分配给云计算节点进行AI模型推断。云计算节点的文案生成部分主要涉及深度学习算法技术，利用预处理得到的信息进行运算并生成文案。并且云计算节点是可拓展的，可以适应高并发的用户请求，也可以保障可靠性。

同时，主服务器上还有用户账号管理系统和数据库，可以支持用户登录和信息同步。服务器上的后台管理平台，可以对上传侵权、违禁内容的用户进行封停和解封处理。

# 实现视图

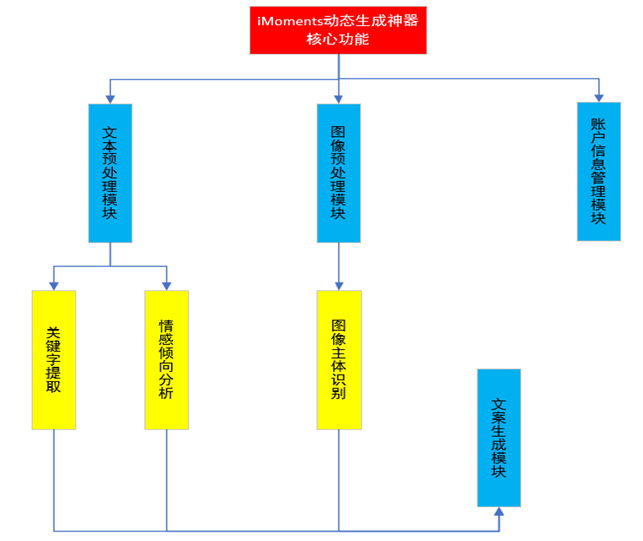


图5 实现视图

# 技术视图

图像通用物体和场景识别： 百度AI开放平台

文字主题和情感倾向提取： 百度AI开放平台

诗歌生成： TensorFlow + WorkingMemory

散文歌词生成： Pytorch + GPT-2

本地藏头诗生成： TensorFlow Lite + LSTM

# 核心算法设计

本地写藏头诗模型采用TensorFlow Lite和LSTM神经网络模型。

云端写散文和歌词采用GPT-2自然语言处理模型。

云端写诗采用Working Memory创作模型。

# 质量属性的设计

每次生成请求的平均响应时间为5秒，最长不超过15秒。

生成诗歌以网络服务的形式注册到主服务进程中，可以快速地扩充算力。

使用Flutter框架可以直接移植iOS设备。