

应用作品展示文档

识秦通

落日古城角，把酒劝
君留。长安路远，
何事风雪敝貂裘。



Fish in troubled waters

TEAM

浑水摸鱼

目录

一、作品简介.....	3
1.1 作品展示形式.....	3
1.1.1 作品展示形式:	3
1.1.2 作品介绍:	3
1.2 设计思路.....	4
二、小程序端介绍.....	5
2.1 用户端展示.....	5
2.1.1 用户登录.....	5
2.1.2 图片上传识别.....	5
2.1.3 探索界面.....	9
2.1.4 最新资讯推送.....	10
2.1.5 用户记录.....	12
2.1.6 兴趣点推荐.....	12
2.2 管理员端展示.....	15
2.2.1 访客统计.....	15
2.2.2 模型详情.....	17
2.2.3 反馈记录.....	19
三、服务器端介绍.....	21
3.1 服务器配置.....	21
3.2 数据库.....	21
3.2.1 用户识别数据库.....	21
3.2.2 用户反馈数据库.....	22
3.2.3 访客记录数据库.....	22
3.2.4 新闻资讯数据库.....	22
四、拓展工作.....	23
4.1 服务器端.....	23
4.2 小程序端.....	23

一、作品简介

1.1 作品展示形式

1.1.1 作品展示形式:

- 1、微信小程序（已完成）;
- 2、Hilens 端侧设备移植（进行中）

1.1.2 作品介绍:

作品名称：识秦通

小程序链接:

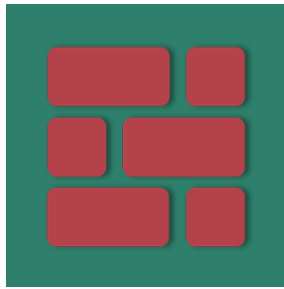
线上版二维码:（由于腾讯对于个人开发者的限制，线上版删除了资讯功能，详情请访问体验版）



体验版二维码:（如无权限，请直接向管理员申请权限）



作品图标:



图标含义：

小程序图标采用“红砖绿瓦”的设计理念，中间的红色砖块代表着西安城墙，背景绿色代表西安钟楼绿瓦，充满了浓浓的古都气息。

1.2 设计思路

根据“华为云杯”2019 人工智能创新应用大赛的举办理念，以“AI 在航天，鸿图华构”为主题，同时通过西安特色景点、美食、民俗、特产、工艺品等展现西安的城市面貌。因此我们设计的作品在完成规定的图片识别任务的同时也可以通过用户上传的图像推广和推荐该类别的详细介绍、POI 信息推荐（用户所在区域周围的感兴趣地点，比如周围的肉夹馍餐馆等）、以及地图指引功能。除此之外，我们在原有识别这个功能的基础上增添了西安周边资讯实时推送功能，让用户能够更加了解不断成长的古都西安。

二、小程序端介绍

2.1 用户端展示

2.1.1 用户登录

为了得到更好的用户体验，我们建议用户初次登录小程序时提供头像昵称信息，方便用户查询自己的识别记录等信息，也有利于统计用户感兴趣的类别，为后期管理员数据分析和资讯定制化推荐提供前期准备（详见图 1a) b)）。

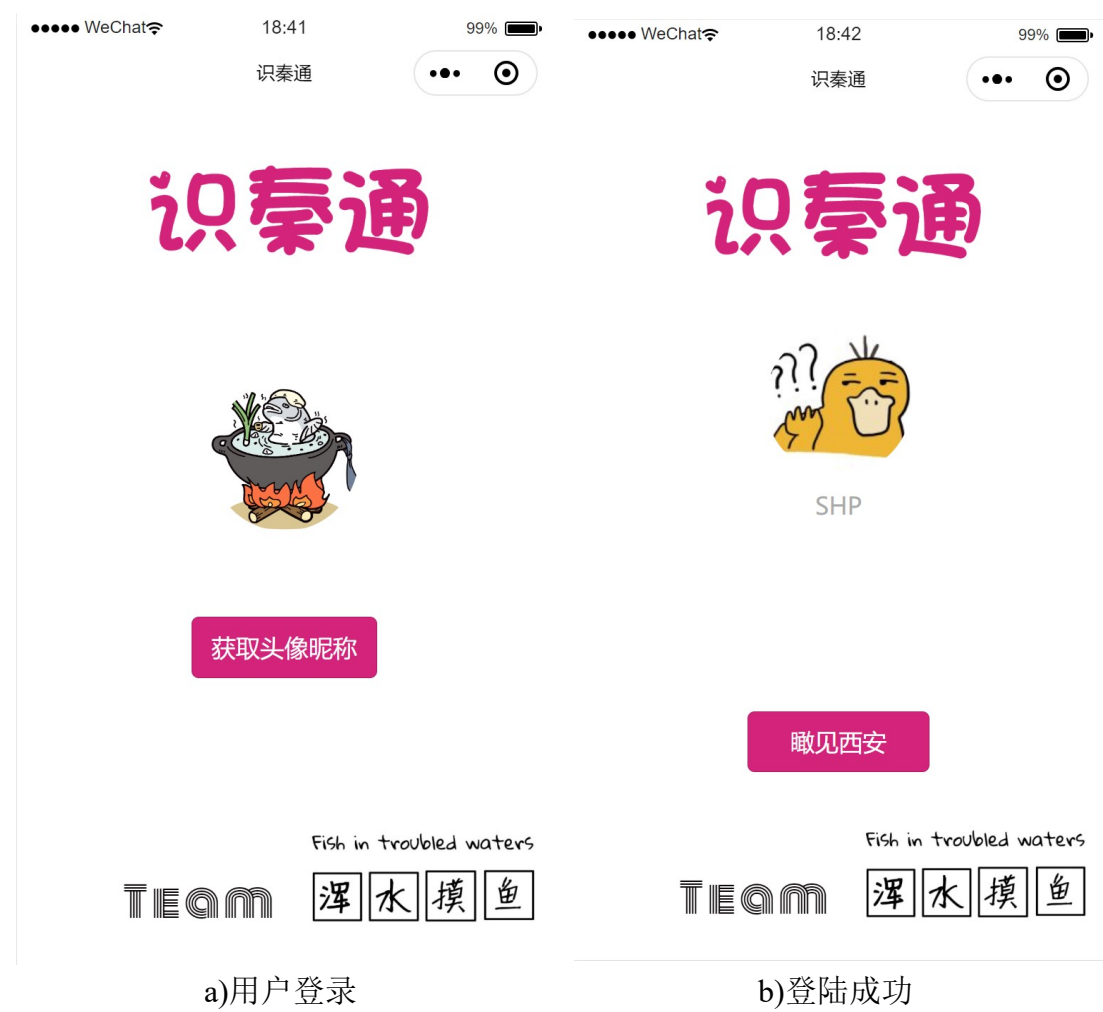


图1. 登录页

2.1.2 图片上传识别

上传图片及识别是小程序的主要功能之一，主要实现的功能是：用户点击上传识别按钮可以从本机相册获取或者手机摄像头中拍摄所要识别的图像，然后服务

器将识别之后的结果反馈给用户，同时展示数据库中与用户搜索图片相似的样例图片和模型对于该图片输出的类激活图(CAM)，方便用户理解模型识别的过程。

用户可以根据识别结果进行反馈，比如分类错误等，也可以点击试试周边寻找附近相关的餐馆、博物馆、建筑遗址等信息。

从图 2c)所示，小程序将为用户展示详细的反馈信息，其中包括：

- 1、用户上传的查询图片
- 2、模型输出的类激活图片
- 3、模型预测概率统计
- 4、基于相似度匹配的图像检索结果
- 5、模型预测类别的百科详细信息
- 6、错误反馈和周边 POI 信息

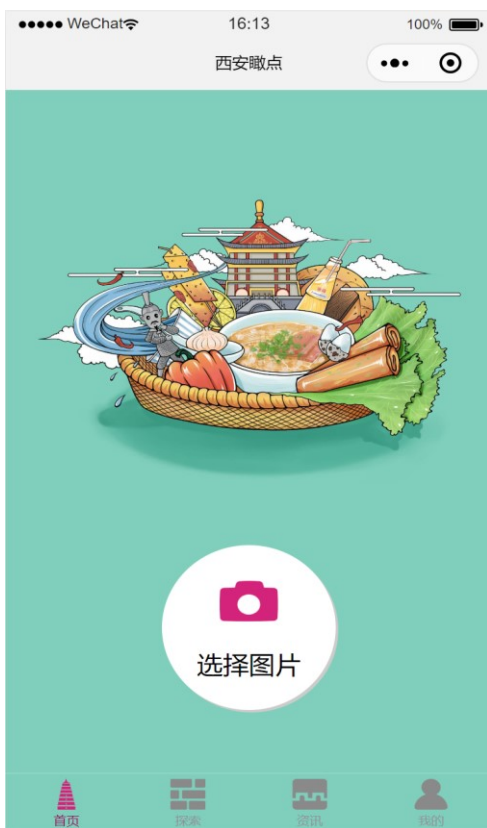
如图 2c)所示，上传图片是用手机拍摄的电脑搜索的百度图片，具有一定的干扰性，但是模型却有很高的置信度认为该图像属于肉夹馍。从反馈的 CAM 激活图我们可以看出，尽管存在电脑边框和周围背景的影响，但是模型只关注了两边存在肉夹馍的地方，忽视了周围的背景干扰，也就是说：模型是通过高响应区域（红色部分）来判断的，因此我们可以认为该模型的识别结果是可信的。

小程序还会展示模型对于上传图片的识别结果，以及 top5 的概率分类情况（详见图 2c)）。为了更好的为用户提供详细的识别过程，我们还提供了样本库（官方数据）中和上传图像最为相似的 top20 图像，以达到用户“以图搜图”的目的（详见图 2d)）。

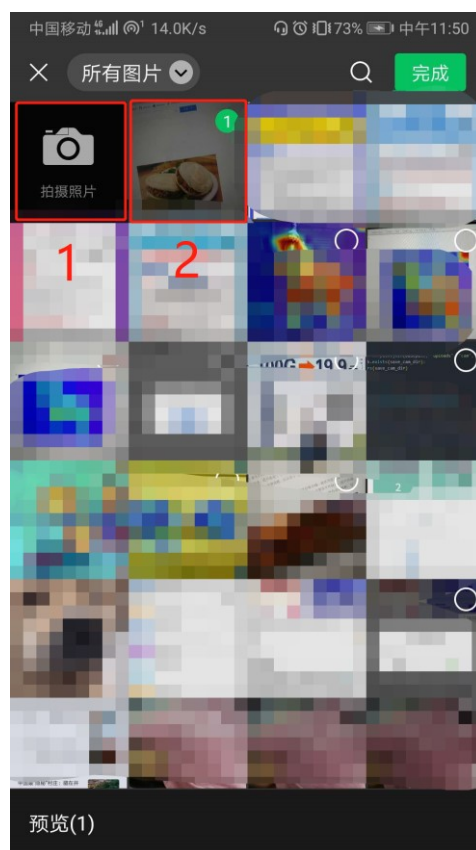
当然，用户可能只是想知道自己识别的图片是什么，并不关系具体的识别过程，那么可以选择百科选项（详见图 2e)），我们为用户提供了该类别详细的百科信息，让用户更好的了解西安。

此外，对于用户上传的图片，模型可能存在预测错误的情况，而且此时用户提供的数据将会是珍贵的难分样本，具有较高的价值，用户可以选择在线反馈操作，见图 2f)。管理员可以在后台中看到用户反馈信息（详见 2.2 管理员端展示），并且可以利用这些信息进一步优化模型，形成良好的正反馈效应。

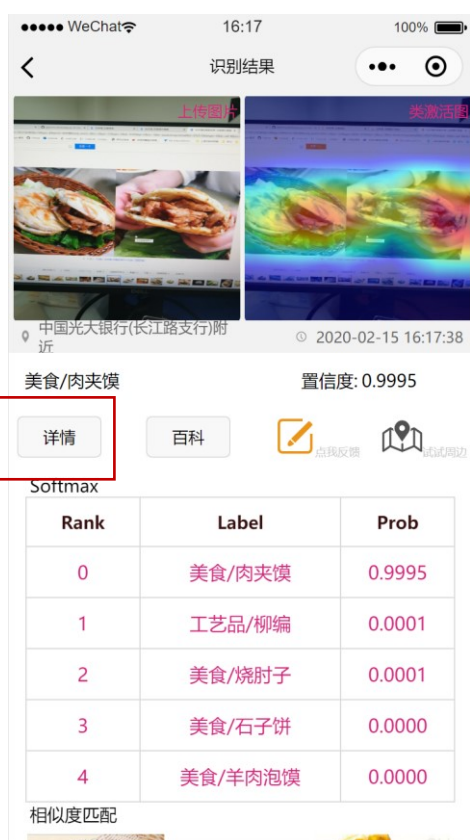
为了进一步的展示西安名片，我们还提供了用户周边与识别内容相关的商铺、建筑遗址、文化遗产等 POI 推荐（详见图 g)），让用户有机会亲身体验西安文化。



a)用户上传



b)选择图片或者拍摄



c)识别结果（详情，top5 概率）

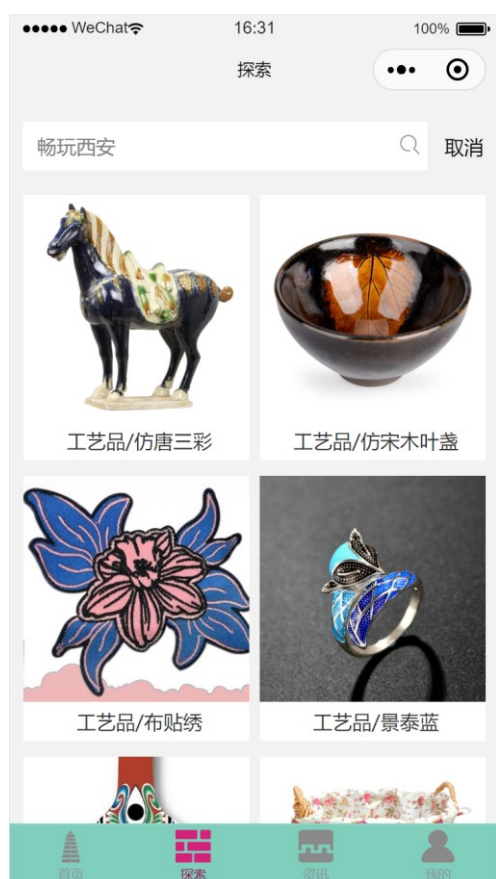


d)识别结果（详情，相似图片检索）

2.1.3 探索界面

对于不了解西安的用户，我们提供了搜索服务，用户可以选择自己搜索感兴趣的类别或者直接在搜索首页中点击类别进行查看。

探索界面载入时会直接显示比赛中规定的 57 类物品列表（见图 3a），列表会通过上拉加载的方式动态加载，加载过程更加流畅。用户可以点击自己感兴趣的类别查看详细的百科信息（见图 3c）。当然，用户也可以自由搜索自己感兴趣的类别，比如：用户搜索美食，那么将会展示和美食相关的类别，更加方便快捷（详见图 3b）。



a)类别探索界面



b)关键字搜索界面



c)百科详情界面

图3. 探索与百科

2.1.4 最新资讯推送

根据小程序设计宗旨，我们同时为用户提供和西安相关的实时资讯，方便用户了解最新的身边事，倾听新时代的西安心跳。

资讯界面分为资讯展示（详见图 4a）和新闻搜索功能（详见图 4b），用户可以点击资讯标题，查看详细的新闻资讯（详见图 4c）。同时，用户可以通过下拉进行新闻资讯刷新，当服务器有新的资讯时，小程序将会展示最新的新闻资讯。资讯会通过上拉加载的方式动态加载，加载过程更加流畅。



a) 资讯推送界面



b) 资讯关键字搜索界面



c) 资讯详情展示

图4. 咨询推送

2.1.5 用户记录

“我的”界面提供了用户的所有查询记录，方便用户查询和反馈，同时提供了滑动删除功能，删除不想要的查询记录。

历史查询结果会通过上拉加载的方式动态加载（详见图 5a），用户可以点击记录查询之前的识别详情，也可以在详情中进行错误反馈（详见图 5b）。同时也可以通过左滑进行记录删除操作。

如果用户是管理员或者在管理员群组里面的话，“我的”页面将会出现一个可以拖动的后台功能，管理员可以通过这个按钮进入管理员界面。

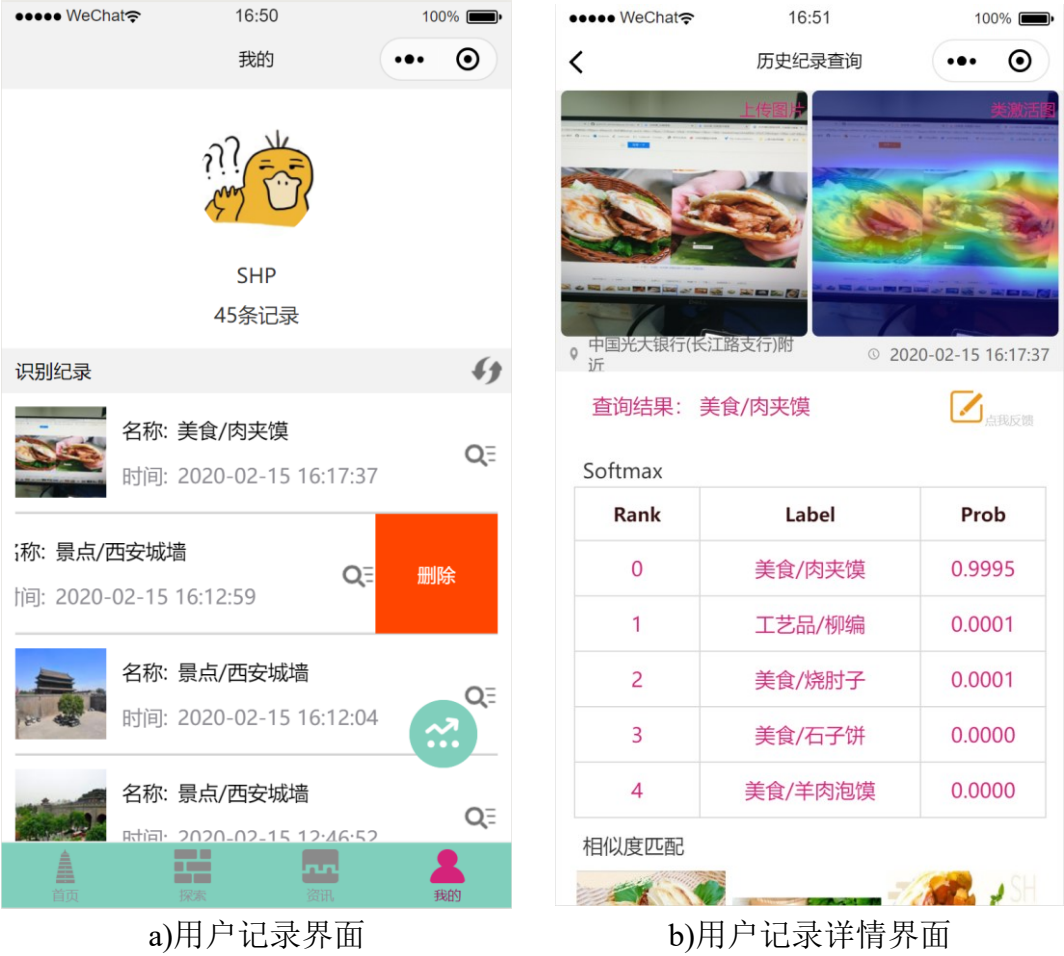


图5. 用户记录

2.1.6 兴趣点推荐

2.1.2 图片上传识别节和 2.1.3 探索界面节，用户识别记录和类别详情百科中，用户可以通过点击“试试附近”按钮搜索用户附近相关的 POI 地点。用户根据 POI 点选择心仪的地点，比如图 6a)用户搜索“凉皮”，小程序将会推荐附近的凉皮商店，用户可以自由选择，并且通过点击“到这去”功能，小程序将会显示不同的出行方式和路线。

1. 驾车方式

驾车页面会提供目标距离、预计耗时、打车费用以及具体的驾车路线（详见图 6b)）。用户可以通过点击“详情”功能查看具体的驾车路线（详见图 6c)）。

2. 步行方式

步行页面会提供目标距离、预计耗时以及具体的行走路线（详见图 6d)）。用户可以通过点击“详情”功能查看具体的驾车路线（详见图 6e)）。

3. 公交方式

公交页面会提供详细的公交路线，包括不同的换乘方式、而且对于公共交通是否停运实时推荐公交方案（详见图 6f)）。

4. 骑行方式

骑行页面会提供目标距离、预计耗时以及具体的骑行路线（详见图 6g)）。用户可以通过点击“详情”功能查看具体的骑行路线（详见图 6h)）。



a)周围兴趣点展示（凉皮）



b)驾车路径规划



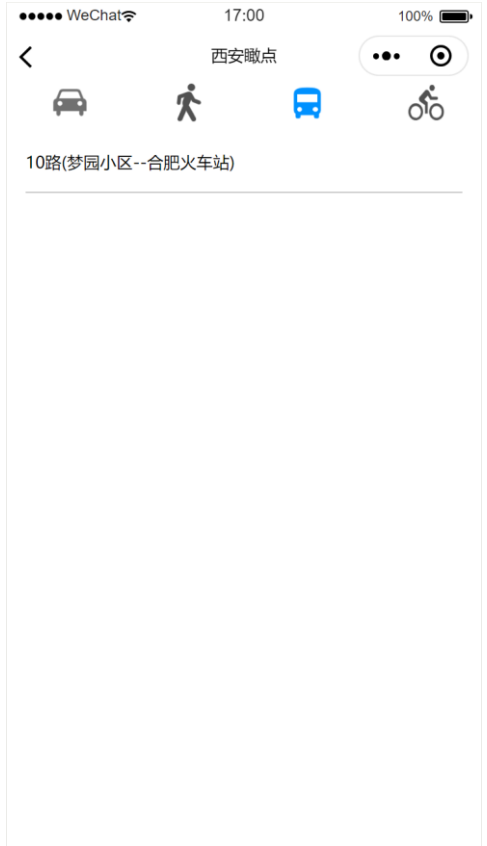
c)驾车路线详情



d)步行路径规划



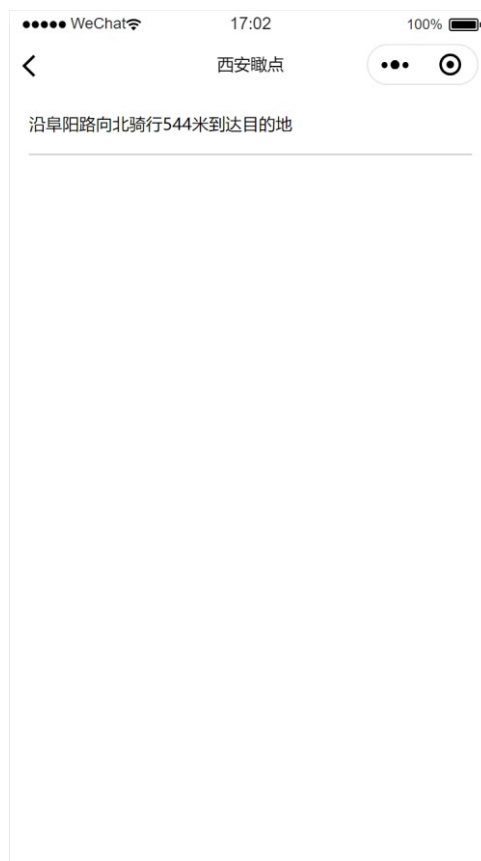
e)步行路线详情



f)公交方式推荐



g) 骑行路径规划



h) 骑行路线详情

图6. POI 推荐路线详情

2.2 管理员端展示

管理员端是为了方便管理员实时了解网站访客请求以及处理用户反馈信息的主要工具，同时增加了识别算法模型的训练详情。管理员可以通过“我的”界面的后台按钮进入管理员界面（详见图 7a）。

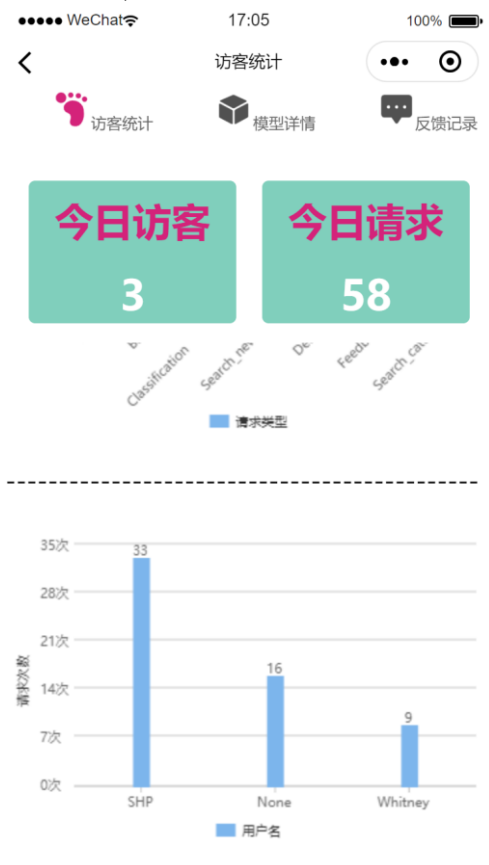
2.2.1 访客统计

访客统计主要展示了有多少用户访问了小程序以及具体的请求形式（详见图 7b) c)），方便管理员了解最新的网络流量和用户感兴趣功能，方便后期用户画像分析以及根据请求类型的不同，进行服务器负载均衡。

访客记录会实时显示今日（截至今日 00: 00）的访客和请求信息，并且绘制最近 24h 用户访问和详细请求趋势。管理员也可以通过时间查询某一时段的访问情况，还可以通过用户的名称查询该用户的查询信息（详见图 7d) e)）。



a) 进入管理员控制台



c) 今日访客统计（请求用户）



b) 今日访客统计（请求类型）



d) 按照用户名查询



e)按照时间查询

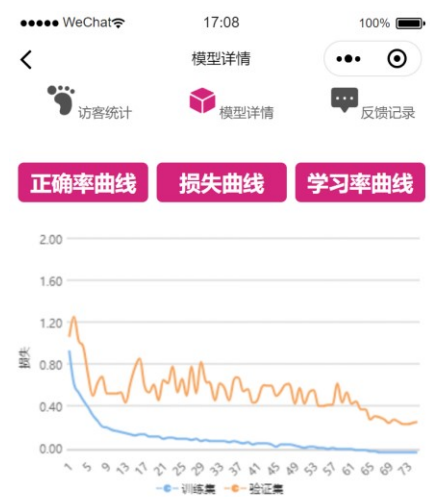
图7. 访客统计页面

2.2.2 模型详情

我们同时提供了模型的训练信息，包括模型训练集和验证集正确率曲线（如图 8a)所示）、损失曲线（如图 8b)所示）、学习率衰减曲线（如图 8c)所示）等。管理员可以动态图标交互，了解模型的训练过程。



a)正确率曲线



b)损失曲线



c)学习率曲线

图8. 模型详情页面

2.2.3 反馈记录

用户反馈对于我们说是非常珍贵的信息，因此我们专门添加了反馈记录模块。管理员可以实时的监控用户反馈信息。

反馈界面展示了今日（截至今日 00:00）的反馈数量和历史反馈数量（详见图 9a)), 图表展示了每日反馈信息变化趋势和用户反馈信息统计（详见图 9b)). 管理员可以按照时间查询某一时间段的用户反馈详情（详见图 9c)), 同时可以按照用户名称查询该用户反馈详情（详见图 9d)), 方便定位用户反馈信息，提取有价值的信息。



a)今日反馈统计（日期）



b)今日反馈统计（用户名）



c)按用户名查询



c)按时间查询

图9. 反馈查询页面

三、服务器端介绍

服务器端主要采用自己的个人 PC，使用 Flask 开启服务，使用 python 处理不同的小程序请求。同时使用 pickle 自己实现了数据库模拟功能。整个服务器系统具有轻量、简洁、灵活、快速的特点。

3.1 服务器配置

由于缺少高性能的云服务器，我们采用了自己的电脑当作服务器，尽量保证电脑服务开启状态。由于我们没有公网 IP，所以必须通过内网穿透之后，小程序才能和本地服务器进行通信。我们采用了花生壳的内网穿透服务。因为免费版用户有带宽和流量的限制，带宽只有 1Mbps，流量每个月 1Gb，因此，小程序加载图片等多媒体数据时可能会有一定的延迟。

微信小程序为了保证用户数据安全，不允许和非 https 验证的域名进行数据交换，而我们之前也没有申请过相应的 https 域名，域名备案更是需要数月之久。我们继续采用了花生壳自带的解决方案，购买了花生壳 https 的一个二级域名。

Flask 默认开启在本地局域网 5000 端口，因此需要一个转发工具，把来自网站的请求转发给 127.0.0.1:5000，同时把反馈信息转发给 https 网站。我们采用了 nginx 作为中转工具。Nginx 是一个高性能的 HTTP 和反向代理 web 服务器，同时也提供了 IMAP/POP3/SMTP 服务。

至此，完成本地服务器环境配置工作。

3.2 数据库

根据业务逻辑，我们设计了四个数据库，分别是用户识别数据库、用户反馈数据库、访客记录数据库和新闻资讯数据库。设置多个数据库的好处一个是简化业务逻辑，数据解耦合。同时方便后期把不同的数据服务配置到不同的服务器上，实现负载均衡和高访问请求响应。

3.2.1 用户识别数据库

用户识别数据库是业务中最为重要的数据库，主要负责存储记录用户的查询操作，记录了不同用户的上传图像、识别结果等，提供了用户查看识别记录、删除识别记录和识别详情的功能。

用户识别数据库名称为‘database.pkl’，使用 pickle 作为存储载体，使用 python 来实现数据库的更新、保存、查询、删除等功能。同时为了防止多用户共同操作或者因为数据库更新和保存冲突，数据库加入了‘上锁’功能，当数据库正在处理某一个需要更新数据库的请求的时候会给当前进程分配一个钥匙，其它数据库访问将会进入排队等待状态，直到锁状态接触，分配给下一个进程钥匙。

3.2.2 用户反馈数据库

用户反馈数据库是主要负责存储记录用户的反馈操作，记录了不同用户的反馈信息等，提供了管理员查看反馈记录的功能，能够更加有效的利用用户反馈的有价值信息。

用户识别数据库名称为 ‘Feedback.pkl’，使用 pickle 作为存储载体，使用 python 来实现数据库的更新、保存、查询等功能，同时加入了数据库访问冲突处理机制。

数据库实现了通过时间段和用户名查询的功能，方便管理员操作。

3.2.3 访客记录数据库

用户反馈数据库是主要负责存储记录用户的请求操作，记录了不同用户的请求情况。提供了管理员查看请求记录的功能，能够更加有效的利用用户请求详情，生成用户画像和协调服务器负载均衡。同时也可以通过用户搜索习惯定制化推荐。

用户识别数据库名称为 ‘query_records.pkl’，使用 pickle 作为存储载体，使用 python 来实现数据库的更新、保存、查询等功能，同时加入了数据库访问冲突处理机制。

数据库实现了通过时间段和用户名查询的功能，方便管理员操作。

3.2.4 新闻资讯数据库

新闻资讯数据库是主要负责存储最新的新闻资讯，实时爬取和西安相关的最新资讯，提供了用户查看最新资讯的功能。

用户识别数据库名称为 ‘News.pkl’，使用 pickle 作为存储载体，使用 python 来实现数据库的更新、保存、查询等功能。

新闻资讯主要通过 python 爬虫实时爬取西安新闻网的资讯 (http://news.xiancn.com/node_2057.htm)，查询更新操作在服务器启动以及用户刷新时触发，做到了实时更新咨询的功能。

因为小程序时属于个人开发者，我们不能简单地使用小程序通过链接的方式访问公网链接，我们需要把新闻内容爬取到本地进行存储。同时，我们选择的新闻资讯网站是 http 域名，新闻里面的图片不能够通过小程序 html 进行解析显示，这就需要先把图片存储在本地服务器，然后使用 https 域名和本地的图片链接替换原始新闻中的图片链接，优化图片在小程序页面的显示体验。

四、拓展工作

4.1 服务器端

1、为了快捷开发，我们没有使用标准的数据库，自己简单的实现了一个小型数据库，结构简单，但是稳定性较差，缺少很多数据库优良特性。后期可以将数据库移植到 MySQL 数据库，提高数据库稳定性和性能；

2、代码结构整理，精炼代码；

3、根据用户请求事件，建立用户画像，为每个用户提供感兴趣内容推荐；

4、扩大资讯爬取范围，建立资讯分类工作。

4.2 小程序端

1、优化交互逻辑，增添管理员数据库管理操作；

2、增加管理员对于用户反馈回应功能；

3、清理冗余代码，精简代码。