



第十四届全国大学生软件创新大赛

文档编号: SWC2021-T20210533-花生队

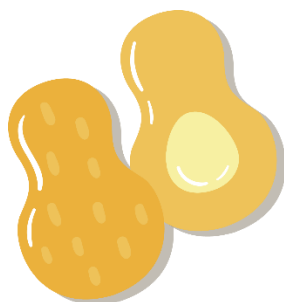


落笔云烟

Magpie's Pen

项目开发文档

Version: V1.2.3



花生队

2021.01.05

All Rights Reserved

目录

1	项目概述	1
1.1	项目背景	1
1.2	项目定位	1
1.2.1	应用场景	1
1.2.2	目标人群	1
1.3	项目方案	2
1.4	项目目标	2
1.5	项目价值	2
2	开发计划	3
2.1	最终呈现形式	3
2.2	主要功能描述	3
2.3	运行环境	3
2.4	验收标准	4
2.5	关键问题	4
2.6	进度安排	4
2.7	开发预算	5
3	可行性分析	6
3.1	技术可行性分析	6
3.2	资源可行性分析	6
3.3	市场可行性分析	8
4	需求分析	10
4.1	数据需求	10
4.1.1	静态数据	10
4.1.2	动态数据	11
4.1.3	数据词典	11
4.1.4	数据采集	12
4.2	功能需求	13
4.2.1	书写图像识别功能模块	14
4.2.2	单字分析与评价功能模块	16
4.2.3	用户个性化服务功能模块	17
4.3	性能需求	19
4.3.1	时间特性	19
4.3.2	适应性	19

4.4	界面需求.....	20
4.5	接口需求.....	20
4.5.1	硬件接口	24
4.5.2	软件接口	25
4.6	其他需求.....	25
5	概要设计.....	26
5.1	处理流程.....	26
5.2	总体结构设计.....	27
5.3	功能设计.....	28
5.4	用户界面设计.....	29
5.5	数据结构设计.....	29
5.6	接口设计.....	35
5.6.1	外部接口	35
5.6.2	内部接口	37
5.7	错误/异常处理设计.....	37
5.7.1	错误/异常输出信息	38
5.7.2	错误/异常处理对策	38
5.8	系统配置策略.....	39
5.9	系统部署方案.....	39
5.10	其他相关技术与方案.....	40
6	数据库设计.....	42
7	手机端侧部署设计.....	43
7.1	手机环境需求.....	43
7.2	调用 AI Unit 能力接口设计（可选）	43
7.3	AI Boost 端侧模型加速设计（可选）	43
8	详细设计.....	44
8.1	**功能模块.....	44
8.1.1	功能描述	47
8.1.2	性能描述	47
8.1.3	输入	47
8.1.4	输出	47
8.1.5	程序逻辑	47
8.1.6	限制条件	48
8.2	**功能模块.....	44

文档修订历史

序号	修订原因	版本号	作者	修订日期	备注
1	创建开发文档	V1.0.0	A	2020.11.15	
2	创建项目概述模块	V1.0.1	B	2020.11.19	
3	创建开发计划模块	V1.0.1	A	2020.11.20	
4	创建可行性分析模块	V1.0.1	D	2020.11.22	
5	更新可行性分析模块	V1.1.0	D	2020.11.23	
6	创建需求分析模块	V1.2.0	C	2020.11.26	
7	更新需求分析模块	V1.2.1	C	2020.11.30	
8	更新模块 1 至模块 4	V1.2.2	ABCD	2020.12.08	
9	更新模块 1 至模块 4	V1.2.3	ABCD	2020.12.30	
10	创建概要设计模块	V2.0.0	A	2021.1.20	
11	完善概要设计模块	V2.0.1	B	2021.1.25	
12	文档校验	V2.0.2	A	2021.1.30	
13	创建数据库设计模块	V2.1.0	C	2021.2.2	
14	完善概要设计模块	V2.1.2	B	2021.2.5	
15	完善数据库设计模块	V2.1.3	C	2021.2.10	
16	完善概要设计模块	V2.1.5	D	2021.2.20	
17	创建详细设计模块	V2.2.0	C	2021.2.25	
18	完善可行性分析	V2.2.0	C	2021.3.1	
19	完善项目概述	V2.2.1	A	2021.3.10	
19	完善开发计划	V2.2.1	A	2021.3.15	
20	完善详细设计模块	V2.3.1	B	2021.3.23	

21	完善需求分析	V2.3.3	C	2021.3.25	
22	完善概述与详细	V2.3.3	D	2021.3.30	
23	更改文章逻辑	V2.4.1	AB	2021.4.3	
24	检查排版	V2.5.1	ABCD	2021.4.9	

1 项目概述

1.1 项目背景

在过去，书法是每个中国文人、学者必备的基本功，是中国最为古老的传统文文化，是中国人高深的文化艺术图腾。常言道：见字如面，字如其人，一手好字，受益终生。一个人的书写水平往往决定了别人对他的第一印象，甚至在不知不觉中影响着一个人的人生轨迹。时至今日，当偶然收到一封来自远方朋友的手写信笺的时候，你仍旧会感受到文字的温度。

随着电脑和移动智能终端的普及，文字书写的使用功能更被极度弱化，在生活节奏加速的今天，“写一手好字”甚至成为一种稀缺的能力。很多人将“练一手好字”的想法添加至待办列表，却苦于不知道如何练习；自己临摹之后缺少老师指导，更不清楚问题出在哪里；自己要发现错误还需要与字帖进行反复对比。于是，多数人因畏惧枯燥乏味的练习、不想付出巨大时间代价而放弃。根据市场调研，目前市面上书法辅助类 App 大多数仅能提供视频教学、字帖或范字索引，并不能与用户产生实质性的有效交互。

针对上述的问题，结合现代的软件设计理念，我们希望应用当前深度学习前沿领域的技术，设计一款能够迅速指出用户书写的不足之处、能给出专业的纠正建议、并帮助用户提高书写综合水平的 APP——落笔云烟。

1.2 项目定位

落笔云烟是一款致力于使用先进的深度学习技术实现现代化、智能化书法练习的安卓应用。

1.2.1 应用场景

1. 任何书写作品后想要评估作品美观程度的场景；
2. 任何想要练习书法，但是缺少专业老师指导的场景；
3. 任何想要改善当前的书写水准，却难以长期坚持临摹练习的场景。

1.2.2 目标人群

1. 任何对改善书写水平、迅速发现自身书写问题有需求的学生党、上班族；
2. 想要学习书法，但身边缺乏专人指导的书法爱好者。

1.3 项目方案

1. 针对字帖临摹方式练字局限性较大的问题，本项目通过图像降噪、图像增强、图像分割等技术来实现用户上传图像中手写汉字识别和对用户的手写作品的内容检测，让用户彻底摆脱字帖的束缚，自由选择练习内容，来实现随写、随拍、随练的使用体验，更适合当代社会的使用场景。
2. 针对用户需要老师指导才能找到错误点的问题，本项目通过提取出的字态特征，应用深度学习技术和自然语言处理相关技术，构造字型评分与字态纠正的若干神经网络，为用户智能匹配最优字体，对用户的手写汉字进行多维度评分并给出最佳的修改建议，来实现让用户迅速发现并改正单字的结构性错误，避免了读帖所带来的效率低下。
3. 针对用户对于迷茫于如何练习的问题，本项目通过对用户的书写历史进行记录并建模分析，构建用户的成长画像和学习曲线，根据用户常错字建立具有相似易错结构的推荐字库，来实现为每个用户设计专属的学习路线，迅速纠错，高效进步。
4. 针对练字过程中的枯燥乏味、难以坚持的问题，本项目通过建立用户激励机制，如趣味闯关、天体系统、练习打卡等模块，以及通过风格迁移实现的个人风格字体生成等特色功能，来实现给用户及时、有效的正反馈，增强用户的自信心和满足感，从而形成良性循环，激励用户持续练习。

1.4 项目目标

落笔云烟旨在通过基于深度学习的图像识别算法与手机 APP 的生活使用场景深度结合的方式，借助于目标检测，深度学习模型，显著性分析，模型梯度可视化等技术，从而解决涵盖从普通用户到专业用户在书法练习中的一系列痛点问题。用户不必再依赖于字帖或书法教师，就可获得专业级的书写指导，从而能够更加轻易地提升自己的书写水准以及书法水平；在此过程中，我们还将通过一系列的用户激励极致降低用户在练习时的枯燥感和所需的时间代价，以达到最佳的练习效果和用户体验。

1.5 项目价值

落笔云烟结合移动终端的强大算力，将学习、练习书法的门槛进一步降低。毫无疑问，在电子终端广泛普及的今天，落笔云烟的出现将具有一定的社会意义：它不仅能为使用者提供全新的练字体验，而其现代化的深度学习算法与中华传统汉字文化的碰撞，也将会为书法领域在 21 世纪的发展带来新思路。可以预见，落笔云烟项目将促进传统书法教学领域的进一步转型升级，在未来有着广阔的市场发展空间，对中国优秀传统文化的与时俱进也将产生长远的意义。

2 开发计划

2.1 最终呈现形式

为了满足广大练字群体对于精细化、个性化辅导的需求，落笔云烟最终以一款利用先进的深度学习技术对用户字体进行实时评分、错误反馈、字态纠正的智能化书法练习辅助软件呈现。

2.2 主要功能描述

核心功能：

1. 书写图像识别：落笔云烟将依据现有的先进深度学习技术，对用户上传的书写图像（包括但不限于字帖、作业等）进行切割并分析，并将结果保存于系统内，为后续服务做准备。
2. 单字分析评价：落笔云烟可将用户单独上传或依据其他功能所得字体进行具体的打分，该功能可进一步识别出影响字体美观的因素，如结构、笔画等，并将对应的修改意见反馈给用户。

辅助功能：

1. 书写推荐：依据本项目对用户字体的薄弱的分析，为其推荐某些特定字体供其练习，以强化用户对该结构的掌握。
2. 风格迁移：为用户生成拥有本人书写风格的字帖。
3. 学习追踪：依据历史数据，项目可为用户提供具体字体风格变化情况以及评分变化曲线，记录用户的每一次成长。
4. 天梯系统：项目为解决现有练字所不得不面对的费事及枯燥，通过设定字体闯关等方式，为用户设定成长系统，变被动练习为主动训练。

2.3 运行环境

表 2.3-1 运行环境表

硬件环境	CPU（Intel）	Memory	Storage	基准频率	RAID
硬件设备 1 （HUAWEI Mate 30）	Kirin990	6GB	128GB	1.25Gbps	1
硬件设备 2 （oppo reno5 pro+）	骁龙 865	12GB	256GB		0
系统环境	Android 10.0				
应用环境	后端框架：Flask				

数据库: MongoDB
应用服务器: python 服务器
前端代理服务器: Nginx
文件共享服务器: dropbox、飞书
任务调度: celery

2.4 验收标准

核心功能模块:

1. 书写图像识别: 对于用户输入的文字图像信息, 在 95% 的情况下, 产品响应时间将不超过 的情况下, 产品响应时间将不超过 2000ms, 在其他情况下不超过 4000ms。保证对单字的切割正确率 $\geq 95\%$ 。
2. 单字分析评价: 对于上传字体评价, 保证模型在美观是具有 95% 以上的准确率, 在修改意见反馈上应保证 90% 以上的准确率; 保证整体响应时间不超过 2000ms。

辅助功能模块:

1. 书写推荐模块: 能够基于用户历史上传字体打分自动训练对应模型, 保证模型具有 95% 以上的推荐正确率。
2. 风格迁移模块: 能够基于用户历史上传字体风格自动训练对应模型, 保证模型具有 95% 以上的推荐正确率。
3. 学习追踪模块: 能够根据具体用户历史数据, 统计出用户具体字体风格变化及分数变化情况, 并以一种用户友好的方式展现, 响应时间小于 1000ms。
4. 天梯系统模块: 能够通过设定合理的分值, 将用户分配至不同的区间, 并可以通过正反馈调动用户练习积极性, 保证赛季时间的合理性。

2.5 关键问题

技术问题

1. 手写字检测时的噪声问题: 在使用目标检测模型前先使用 CycleGAN 去除大部分的图像噪声, 之后使用调优后的目标检测模型检测单个手写字。
2. 用户书写的评价问题: 书法是一门即包含艺术性也需兼顾结构化的领域, 因此书法的评价上不可使用单一的方式评价。本项目通过检测出的汉字骨架, 结合模版汉字的骨架, 从是否缺少笔画、各笔画的差异、结构的问题三个方面进行评分。
3. 用户书写纠正: 改模块分为汉字骨架识别与汉字字形纠正两部分, 汉字骨架识别通过识别出汉字的笔画信息提供给汉字字形纠正部分作为基础信息。汉字纠正部分根据识别出的汉字骨架, 通过计算用户写的汉字与模版

汉字的各个笔画的区别，判断各个笔画的书写情况。之后通过数据库中预先建模好的汉字的结构信息，匹配该汉字存在的书写缺陷及指导建议反馈给用户。

4. 项目数据集获取问题

目前已经具备的开放数据集：

- 1) 中科大手写汉字数据集 CASIA-HWDB
- 2) 哈工大手写识别数据集 HIT-OR3C
- 3) 北邮脱机手写汉字数据集 HCL2000

除此之外，项目团队在上述数据集基础上自行标注了小规模 Peanuts-HWDB 数据集，用于项目中文本评分和字态纠正神经网络的训练测试；目前已经使用其中大约 5% 的可用数据训练模型 demo，已证实可行。

2.6 进度安排

表 2.6-1 进度安排表

项目重要节点	完成时间
组建团队	2020.10.17
前期调研与分析	2020.10.26
确定项目主题与创意	2020.11.07
确定成员分工	2020.11.09
数据集收集	2020.12.05
前端代码编写	2020.12.30
后端代码编写	2021.12.31
模型代码实现	2021.02.20
模型优化与应用测试	2021.03.11
美工优化	2021.03.23
相关代码融合	2021.03.25
代码优化	2021.04.15
模型对接	2021.05.01

2.7 开发预算

表 2.7-1 开发预算表

项目名称	预算 (CNY)
硬件设备 1 (HUAWEI Mate 30)	3999
硬件设备 2 (oppo reno5 pro+)	4499
阿里云 GPU 轻量型服务器 ecs.vgn5i-m1.large	576.5/m

3 可行性分析

3.1 技术可行性分析

1. 手写字检测

手写字检测是从用户输入的书写图像中检测出所有单个的字，并返回单个字的边框信息。这可以看作目标检测问题。目前主流的目标检测算法皆已成熟，本项目使用基于 Faster R-CNN 的模型架构，实现可以从输入图像中检测手写字的模型。

2. 整体书写评价

整体书写评价涉及到文字提取与整齐度计算两个部分，文字提取可以使用目标检测方式完成，技术性一般，整齐度计算主要使用数字图像处理的方法完成，前者使用手写字检测问题中实现的模型来完成提取，后者可以通过数字图像处理的方法计算文本的整齐度，两者实现都相对简单。

3. 汉字骨架识别

汉字骨架识别的目的是检测出用户手写的汉字的各个笔画，我们从 CMU 提出的用于人体姿态识别的 OpenPose 框架中得到启发，设计了可以检测所有汉字骨架的模型，模型的流程见下图。

我们提出的汉字骨架识别方法使用 PAFs (Part Affinity Fields) 来编码汉字书写图像中的笔画方向信息，通过预测关键点位置的置信图来得到不同笔画的关键点的位置。得到关键点的置信图和笔画的 PAFs 后，便可以通过贪心算法来计算不同连接的分数，之后通过识别出的汉字信息，构造可能的汉字笔画的集合，通过计算不同集合的分数得到可能性最大的连接集合，将集合中的关键点与相连即可得到汉字的骨架结构。

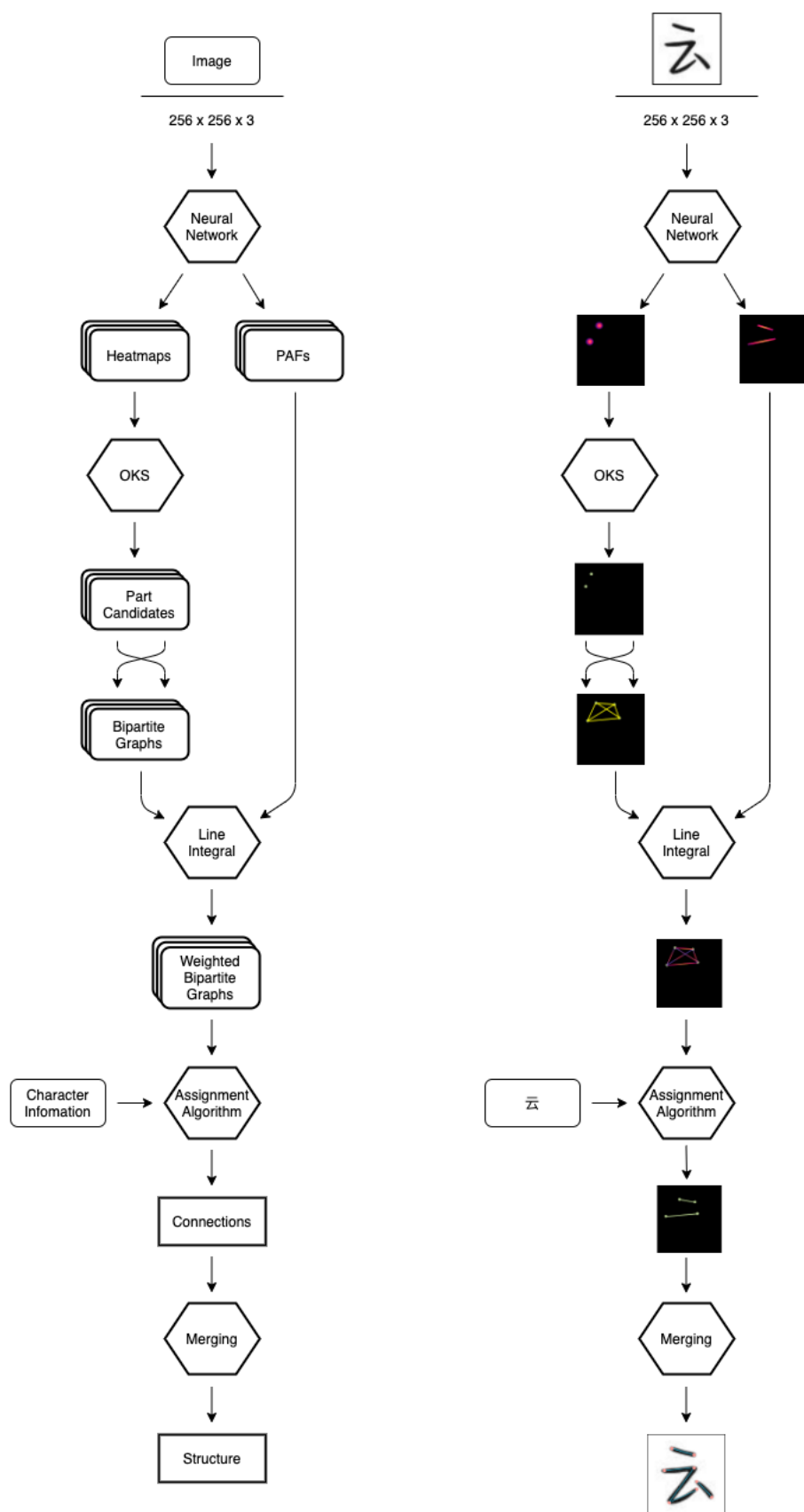


图 3.1.1 汉字骨架识别模型

4. 书写汉字纠正

书写汉字纠正共分为三个模块，首先是计算输入汉字图片与模版汉字图片之间的最适合匹配位置，并生成这两者的蒙版，将其反馈给前端供用户查看自己书写的汉字与标准字之间的差异所在。其次是通过对比检测出的输入汉字的骨架与标准字之间骨架的差异，依据不同的尺度计算这两者的不同，并将问题反馈给用户。最后是根据识别出的汉字信息，在我们事先构建的专家系统中匹配该汉字对应的一系列需要注意的书写建议，根据检测出的汉字骨架判断是否存在相应的问题，最后将问题与建议反馈给用户。

3.2 资源可行性分析

数据资源

1. 中科大手写汉字数据集 CASIA-HWDB
2. 哈工大手写识别数据集 HIT-OR3C
3. 北邮脱机手写汉字数据集 HCL2000
4. 项目团队自研数据集 Peanuts-HWDB

计算资源

目前项目拥有实验室服务器资源使用权。目前可使用的计算资源如下：

1. 4*GeForce RTX 3090
2. 8*GeForce GTX 1080
3. 4*GeForce RTX 2080
4. 9*GeForce GTX Titan Z

3.3 市场可行性分析

如何实现写字由不好看变为好看的转变一直以来都是一个非常大的挑战，此项目灵感也来源于我们团队成员在疫情时期练字的感悟。在经历了传统的练字练习之后，我们思考如何才能优化其中体验感差的部分，真正地提升大家的效率。

于是借软件大赛的契机，我们展开了对于这一个领域与发展现状的思考，并且思考如何加持人工智能，如何加持端云结合，精准发掘这一领域的上升空间。经过一段时间的需求分析和市场调研，我们发现此款应用在市场上是完全可行的，接下来是我们的分析思路与分析结果。

1. 目前大家练字的手段主要有两种，一种是请私教，包括一些辅导机构；另一种是自己查阅资料买字帖，自己一个人练习。

- 大家练字的目的也有两种，一种是专业性比较强的，想要在书法方面有所造诣；另一种是从实用角度上，想让字的整体性变得“好看”即可。本应用针对的主要方向为第二种，第一种为辅。

对于大家练字的第一种手段——请私教，更适用于第一种目的，即想要在专业性上有所提升，而对于想要从使用角度上提升的人群，更多使用的方法是自己找教程练习字帖。

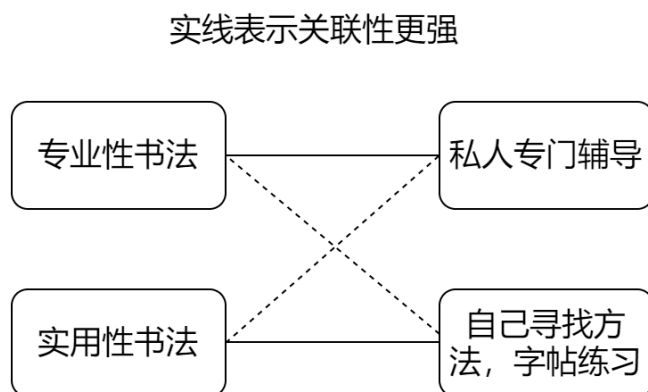


图 3.3.1 目前市场目的与手段关联性

而对于练习实用性书法的人群，现在的方式都有些弊端。

- 如果选择私人专门辅导，那么付出的时间、金钱、精力都要远远高于个人练习，对于练习实用性书法来说没有必要。
- 如果选择个人练习，那么常见的解决思路有两种，一种是寻找网络上前辈的经验分享，比如知乎、豆瓣等，一种是自己寻找一种风格的字体，从头开始练。

对于第一种借鉴别人经验分享，在某些情况下是有一定效用的，但是往往这些方法都比较个性化，受幸存者偏差的影响，能够分享出来的方法一定是适合练字者本身的，但不一定适合其他人，于是就导致了市场上的方法参差不齐，我们一眼望过去也不知道如何选择，如何有效。

第二种方法是基于第一种，但是超于第一种的。我们知道练字需要持之以恒，不能急于求成，跟着字帖练字会比第一种靠谱很多。但是这种方法也暴露出很多弊端：(1)是否从头开始练习，效果固然会更好，但是代价也是更难坚持下去；(2)没有即时的反馈，即使能够坚持下去，但如果不进入积极思考寻求改变，会发现只是在练但进步却不明显，离开字帖发现还是老样子。练字本身就是挑战自我、突破自我的过程，此时如果有辅助工具相助，一定会更加有成就感与乐趣。

以上简单分析了一下目前练字的解决方案，而 APP 等电子辅助产品与设备大多是应用在自己的练习临摹阶段。于是我们对于目前市面上的练字类 APP 进行分析，分成了查询类与智能类两类。接下来依次进行分析。

对于查询类 APP，目前做这一方面的 APP 有很多，比如：不厌书法、以观书法等（更详细的目录可见竞品分析），常见的思路是录入碑帖中的字在他们的数据库中，然后用户想查哪个字就去看看然后自己再进行临摹，对于临摹之后，

大多数 APP 做的工作是进行两个字的比对，一般只是罗列出来让你自己比对，做的比较好的会是两个字做一个重叠图层，更方便观看。这一类 APP 不管是查询还是临摹对比，最终还是回归到了自己练习自己纠错的阶段，核心功能是查询功能。

对于智能类 APP，比较典型的是方正习字。这款 APP 做的工作是选择特定的字之后可以在纸上书写然后拍照，自动分析跟你说明两个地方：某个笔画过长过短、某个笔画倾斜角度不对。经过一段时间的体验我们发现有以下问题，一是这款 APP 的智能分析只针对软笔书法的大约 200 个字，对于硬笔字和其他大部分软笔只有分享功能没有测评功能，二是它对于字的测评比较简单，基本每次都是一个模板替换其中笔画不一致的地方，提供的作用有限。

于是我们结合用户的需求与现在市场上不能解决的问题，开发出自己的 APP。我们旨在填补市场上的空隙，在 APP 的人工智能提示下能够引导用户动脑，加速孵化过程，在实用性上拔得头筹。

我们深知，练字不能急于求成，练字的核心在于改变，贯穿着内心态度到实际执笔书写。我们选择作为帮助用户改变的使者，学会分析，善于思维的加速。基于足够理性且逻辑化的大框架，遵守相对规则进行系统化训练，才能取得高效练字效果。我们要成为用户在前进道路上的良师益友，从而高质量的练字，快而不浮躁，静而寻求改变，在这一领域的市场中绽放光彩。

4 需求分析

4.1 数据需求

4.1.1 静态数据

静态数据，又称“横截面数据”，是指在运行过程中主要作为控制或参考用的数据，它们在很长的一段时间内不会变化，一般不随运行而变。在此，我们用到的静态数据主要分为以下两方面：

表 4.1.1-1 静态数据表

	名称	设置值	定义	格式	类型
服务器操作部分	模型地址	/model_path/..	用于保存训练后的模型的地址	String	String
	sql	'select * from ...'	数据库操作的固有命令，用于直接处理信息	String	String
	finalPath	/image/./..	用于保存用户上传的图像信息	String	String
				

4.1.2 动态数据

表 4.1.2-1 动态数据表

	名称	定义	格式	类型
服务器操作部分	用户个人信息	用于保存用户个人设置及隐私信息	struct	struct
	服务器流量	用于保存当日访问人数	integer	integer
	用户字体分数	用于保存用户拍照字体历史得分数据	double	double
	用户天梯分数	用于保存用户天梯得分数据	String	String
			

4.1.3 数据词典

表 4.1.3-1 用户信息表

字段	描述	类型	允许空值	默认值
id	用户 id	int(10) UNSIGNED	No	
name	用户名	varchar(255)	No	
email	用户邮箱	varchar(255)	No	
email_verified_at	邮箱验证时间	timestamp	Yes	
password	用户密码	varchar(255)	No	
remember_token	找回密码令牌	varchar(100)	Yes	
created_at	创建时间	timestamp	Yes	
updated_at	更新时间	timestamp	Yes	
avatar_url	头像地址	varchar(255)	No	/avatar/default_avatar.png'
phone	手机号	varchar(255)	No	
birthday	生日	date	No	1901-01-01'
gender	性别	tinyint(1)	No	0
structure	用户字体布局得分	int(11)	Yes	

表 4.1.3-2 用户上传字体信息表

字段	描述	类型	允许空值	默认值
word_id	字体 id	int(10) UNSIGNED	No	
user_id	对应用户 id	int(10) UNSIGNED	No	
user_img_path	用户字体图像存储	varchar(255)	No	
grade	用户得分	double	No	0

date	字体得分日期	date	No	1901-01-01'
YB	运笔得分	int(11)	No	0
DX	大小得分	int(11)	No	0
JG	结构得分	int(11)	No	0
ZX	中心得分	int(11)	No	0
JD	角度得分	int(11)	No	0
style	该字体风格	int(11)	Yes	
img_path	标准字体图像 存储	varchar(255)	No	

表 4.1.3-3 天梯信息表

字段	描述	类型	允许空值	默认值
user_id	用户 id	int(10) UNSIGNED	No	
time	赛季时间	varchar(255)	No	
point	用户该赛季点数	int(11)	No	1200

表 4.1.3-4 找回密码表

字段	描述	类型	允许空值	默认值
email	验证的邮箱	varchar(255)	No	
token	找回密码令牌	varchar(255)	No	
created_at	创建时间	timestamp	Yes	

4.1.4 数据采集

数据采集工作分为公开数据采集和团队自研数据库采集。

目前已经采集到的公开数据有：

1. 中科大手写汉字数据集 CASIA-HWDB
2. 哈工大手写识别数据集 HIT-OR3C
3. 北邮脱机手写汉字数据集 HCL2000

团队自研数据库目前已经初具规模，为此团队专门构建了简易标注客户端，正在通过网络爬虫和有偿人工标注等方式进行进一步扩充。

4.2 功能需求

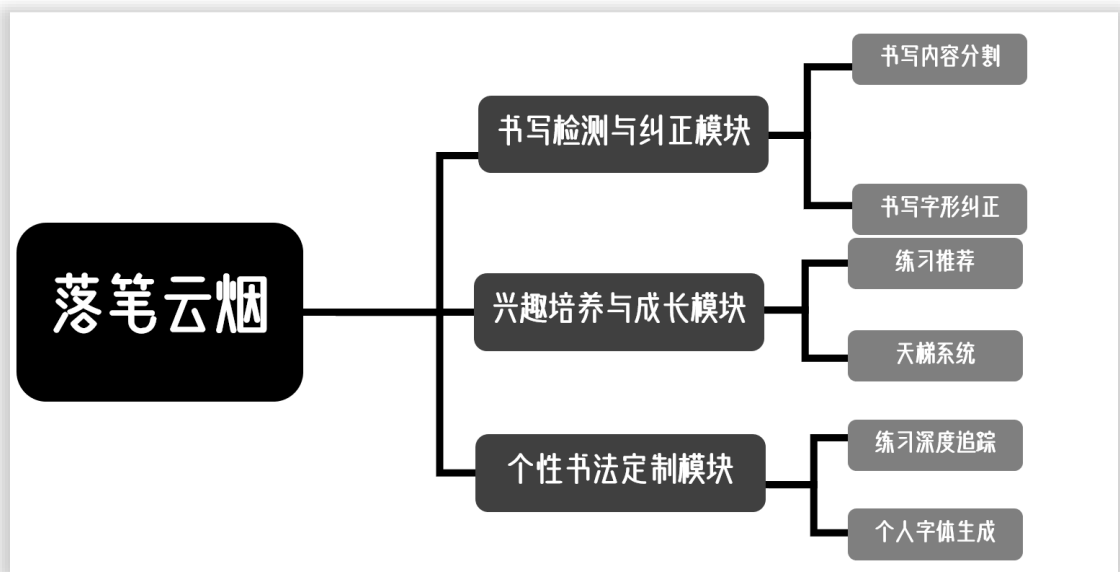


图 4.2.1 功能模块结构图

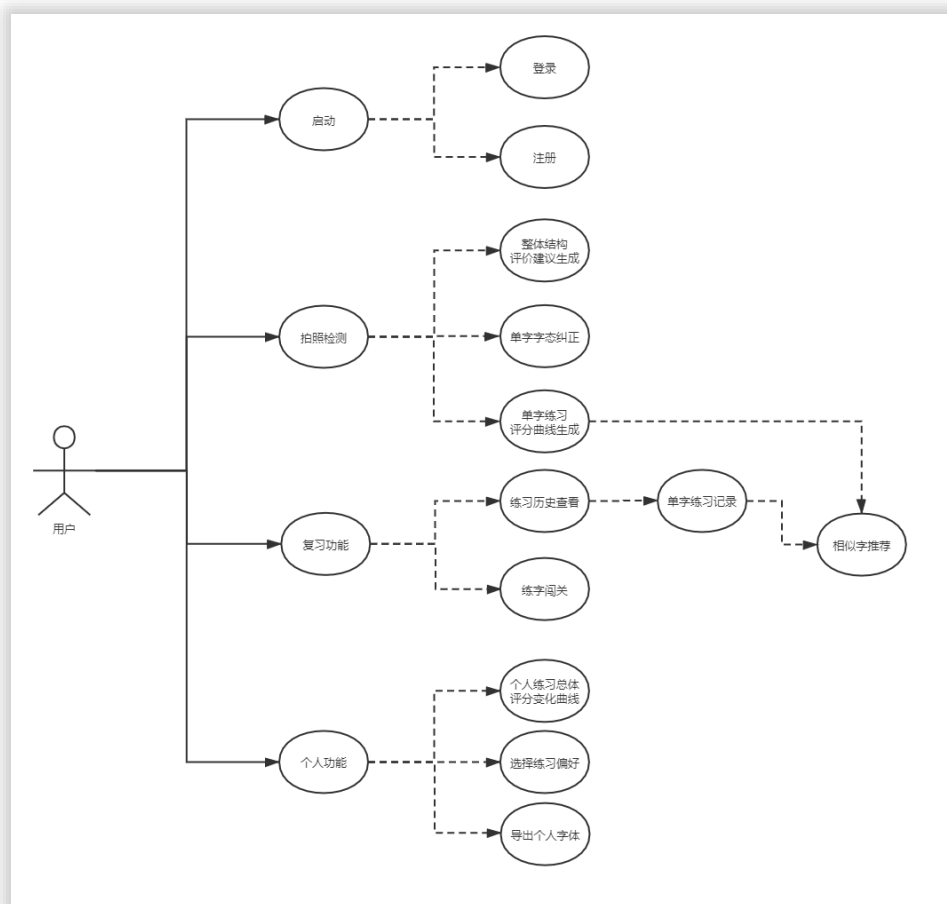


图 4.2.2 用例图

4.2.1 书写图像识别功能模块

表 4.2-1 书写图像识别功能模块描述

功能模块	功能	功能描述	优先级
书写图像识别模块	整体图像预处理	整体图像的处理，去除图像中的噪声，并做一定程度的图像增强。	6
	整体书写情况评价	建立多维度评价模型，结合处理后的图像与识别出的单个字体进行多元化的总体情况反馈。	7
	单字识别与分割	识别出图像中手写的字，忽略非手写的字，例如图像中包含的印刷字等。	6
	书写报告生成	综合整体书写评价，使用自然语言处理技术生成使用户易于理解的书写报告。	4

表 4.2-2 整体图像预处理用例规约

用例名称	整体图像预处理
功能简述	对用户上传的图像进行降噪与图像增强，为后续功能能做预处理。
用例编号	LBYY.001
执行者	LocalServer
前置条件	① 手机拍摄的 jpg 照片 ② 照片小于 20MB ③ 有效文字分辨率至少为 50px*50px ④ 噪点低于阈值
后置条件	生成对于后续功能来说质量和成功率更高的 jpg 文件
涉众利益	① 整体书写评价：行与行之间有严格区分度，字与字之间有较高区分度 ② 单字分析与评价模块：字与字之间有严格区分度
基本路径	① 用户点击拍照评价模式 ② 选择现场拍照或者相册照片 ③ APP 进行处理图片 ④ 如果图片合格，进入下一个功能模块
扩展路径	① 图片过大给予提示并退出 ② 分析过程中图片不满足分辨率分析要求给予提示并退出 ③ 图片干扰与噪点过多提示并退出
字段列表	图片存储格式、图片名称、用户 token
设计规则	抽离为单独模块，为后续模块提供接口
未解决的问题	不支持过于凌乱的字体和超过识别容量的纸张大小
备注	是其他步骤的基础，要权衡效果与鲁棒性；还要注意对用户的提醒方式要友善。

表 4.2-3 整体书写情况评价用例规约

用例名称	整体书写情况评价
功能简述	经过上表的图像预处理后，建立多维度评价模型，生成此次评价数据向量。
用例编号	LBYY.002
执行者	MainServer
前置条件	① 整体图像预处理生成的新图像文件 ② 下文所描述的单字书写数据
后置条件	生成与存储针对此图像数据的评价数据向量。
涉众利益	① 正确生成用户的此次评价数据 ② 为报告生成提供准确的数据输入
基本路径	在图像预处理完成后进行
扩展路径	无
字段列表	用户 token、结果 text 与变量存储
设计规则	抽离为单独模块，能够独立实现对用户的评价输出
未解决的问题	数据量需要降维，做精度的妥协，为了后续给用户的报告生成提供足够快速的体验
备注	仅仅是做数据处理，不生成报告

表 4.2-4 单字识别与分割用例规约

用例名称	单字识别与分割
功能简述	识别出图像中手写的字，减少干扰因素的影响
用例编号	LBYY.003
执行者	MainServer
前置条件	整体图像预处理生成的新图像文件
后置条件	将单字各项数据存储到本次评价的数据结构中，为第二模块单字评价模块做高质量的输入
涉众利益	第二模块单字评价算法需要足够质量和预处理后的单字模型
基本路径	在图像预处理完成后进行
扩展路径	识别率过低则停止
字段列表	单字数据结构存储
设计规则	抽离为单独模块、标准算法模型
未解决的问题	受算力支持，需要妥协和识别率，做时间与效果的权衡
备注	此用例优先级高，算法方面比较有难度

表 4.2-5 书写报告生成用例规约

用例名称	书写报告生成
功能简述	综合整体书写评价，使用自然语言处理技术生成使用户易于理解的书写报告。
用例编号	LBYY.004
执行者	MainServer

前置条件	① 用户画像数据 ② 整体书写评价所给的数据向量
后置条件	将报告结果输出一份，存储一份与用户相关联的历史记录
涉众利益	① 用户：能够正确显示自己的不足、有建设性建议 ② 用户模块：生成有用的可以完善用户画像的数据，包括评分
基本路径	① 得到书写评价处理的数据 ② 生成报告
扩展路径	生成时间过长给予用户提示是否继续等待，如果超过一定时间强行终止
字段列表	报告生成输出接口对象
设计规则	抽离为单独模块、标准算法模型
未解决的问题	受算力支持，需要妥协时间与效果的权衡
备注	NLP 处理，避免反馈模板化，导致对用户的无效反馈

4.2.2 单字分析与评价功能模块

表 4.2-6 单字分析与评价功能模块描述

功能模块	功能	功能描述	优先级
单字分析与评价模块	单字图像处理	针对单字进行精细化处理，包括处理走样、分辨率等。去除字体图像的噪声，并做一定的图像增强，方便后续流程的处理。	7
	单字全局书写缺陷检测	需要识别出用户书写的字体中的结构问题与总体问题，并以易于理解和接受的方式反馈给用户。	8
	单字局部书写缺陷检测	需要识别出用户书写的字体中可能存在的局部书写缺陷，例如部分笔画的问题，并将结果以易于理解的方式反馈给用户。	7
	单字评分	对用户书写的字体给出合理的评价，用户书写的字体可能有自己的风格，需要一个灵活的模型给出合理的评分。	6

表 4.2-7 单字图像处理用例规约

用例名称	单字图像处理
功能简述	对单字图像进行进一步处理，去除整体噪声，进行局部增强等。
用例编号	LBYY.005
执行者	LocalServer
前置条件	① 用户提供正确的文字图像 ② 单字图像分割完毕
后置条件	输入图像以灰度矩阵形式传入后续模块
涉众利益	① 用户：可以通过直接拍照的方式上传单字信息 ② 单字全局与局部书写缺陷检测模块：获得合适的输入字段

基本路径	用户用过拍照或从相册选择图片的方式上传图像
扩展路径	若图像中不包含分割好的汉字，则提醒用户重新上传
字段列表	图像是否合法，灰度矩阵
设计规则	预处理模块
未解决的问题	由于算法限制，不可避免导致有噪声点的直线变弯以及笔划交叉或分叉位置产生少许畸变的问题
备注	该部分算法基于笔画趋势分析的二值化方法，更适应使用场景。

表 4.2-10 单字评分用例规约

用例名称	单字评分
功能简述	对于用户选择的单字进行横向综合评分，并给出指导性建议
用例编号	LBYY.008
执行者	MainServer
前置条件	① 整体、单字图像处理完毕 ② 用户选择需要进行检测的单字 ③ 单字局部于全局检测进行完毕 ④ 整体书写报告生成完毕
后置条件	① 反馈结果数据给用户 ② 结果将记录在用户的书写历史数据库中
涉众利益	用户希望获得单字的综合评分
基本路径	① 汉字检测，汉字骨架识别后完成后进行 ② 反馈处理数据给前端用于书写报告生成动态曲线
扩展路径	① 无法识别出文字时提醒用户重新上传图像 ② 进行该字的相似性易错推荐
字段列表	单字美观度综合性评分
设计规则	单独模块
未解决的问题	由于美观度含有较多的主观因素影响，评分模型依赖于骨架检测模型，目前骨架检测数据集尚未覆盖所有汉字。
备注	基于骨架检测结果运行汉字评分

4.2.3 用户个性化服务功能模块

表 4.2-11 用户个性化服务模块描述

功能模块	功能	功能描述	优先级
用户个性化服务模块	字体风格迁移	根据书写特征生成用户专属字体文件，以用户的历史书写的字体风格数据为基础，迁移生成个人字体库。	3
	个性化练习推荐	根据书写历史推荐练习内容及相似字体，需要根据用户之间以及单字的相似度来动态计算推荐群及信任子群，同时需要考虑解决冷启动问题。	6

	练字趣味闯关	使用前述功能接口组合, 设计闯关系统, 增强 APP 的趣味性。	1
	天梯等级系统	设计适合于此项目的等级系统, 目的是给予用户正反馈和一定的激励。	2

表 4.2-12 字体风格迁移用例规约

用例名称	单字评分
功能简述	根据用户书写历史数据迁移生成用户的个人风格字体文件
用例编号	LBYY.009
执行者	MainServer
前置条件	用户上传处理超过 N 个汉字
后置条件	通过生成迁移风格字体文件, 存储在用户存储器中或生成分享链接
涉众利益	用户希望获得带有个人风格的手写字体
基本路径	① 用户使用 App 处理一定数量的个人手写文字 ② 根据用户的手写历史数据, 基于基本文字库进行汉字字体风格的迁移 ③ 生成字体文件并保存再手机存储器或生成文件分享链接
扩展路径	① 累计历史数据不足以进行风格迁移时, 提醒用户 ② 存储空间不足时, 提醒用户
字段列表	迁移后生成的个人风格字体文件
设计规则	单独模块
未解决的问题	由于风格迁移模型需要一定的数据作为输入, 因此在用户使用 App 次数较少时难以获取到足够数据支撑迁移。
备注	该部分算法基于 SelectorGAN 等模型, 适用于汉字生成的使用场景

表 4.2-13 个性化练习推荐用例规约

用例名称	单字评分
功能简述	根据书写历史推荐练习内容及相似字体供用户进行练习
用例编号	LBYY.010
执行者	MainServer
前置条件	① 用户上传处理过一定数量的汉字 ② 用户上传处理汉字数量不足时采用冷启动方法
后置条件	将相似度排名前五的汉字加入到推荐列表
涉众利益	用户希望找到相似易错字进行针对性练习
基本路径	① 用户查看单字检测历史记录时进行推荐 ② 用户进行单字检测完毕时进行推荐
扩展路径	无
字段列表	相似度排名前五的汉字编号
设计规则	单独模块
未解决的问	受算法制约, 用户使用 App 次数较少时冷启动的推荐效果相对较

题	差。
备注	该部分基于不确定近邻的协同过滤算法, 实际使用效果应当是“越用越准”

表 4.2-14 天梯等级系统用例规约

用例名称	天梯等级系统
功能简述	设计适合于此项目的等级系统, 目的是给予用户正反馈和一定的激励。
用例编号	LBYY.012
执行者	MainServer
前置条件	用户所有数据
后置条件	生成等级与段位
涉众利益	用户希望看到自己的水平如何和是否有长进
基本路径	个人中心后台计算, 实时显示
扩展路径	初始为 1 级, 暂无满级
字段列表	等级字段
设计规则	单独模块
未解决的问题	无
备注	从用户的角度出发, 围绕能够给用户起激励作用而设计, 前端需要做得美观

4.3 性能需求

4.3.1 时间特性

表 4.3.1-1 响应时间表

首页	/index	少于 30ms
拍照评字	/take_word	少于 100ms
发现	/discover	少于 30ms
复习	/review	少于 100ms
个人报告	/report	少于 100ms
个人通知	/notification	少于 50ms

4.3.2 适应性

落笔云烟应该具有适应能力包括以下几个方面:

1.操作方式: 落笔云烟的用户群体主要分两大类, 一类是时间较为零散的学生党、上班族, 落笔云烟的操作方式需要尽可能地简单, 能够为其快速指出字体不足之处。另一类是书法的爱好者, 需要依据其个性化的字体优化方向, 给予较为专业的练习推荐。

2.运行环境：落笔云烟最终呈现形式为一款 Android 移动端的应用，因此需要在各类基于 Android 的移动设备上运行无误。

4.4 界面需求

高保真原型设计稿(使用 Figma 制作)

1) 首页设计：

将主要功能直接呈现给用户；用户可以选择开始识别进行拍照，或者从相册中选择图片进行上传，或者选择相关文章进行浏览。



图 4.4-1 首页设计图

2) 文章详情：

展示从首页进入的文章详情，文章类型为展现汉字历史、排版文化的相关知识，偏向于娱乐性。后续有计划使其直接能够成为用户练习的文章，使用户可以直接临摹这个练习，方便用户找练字资源。

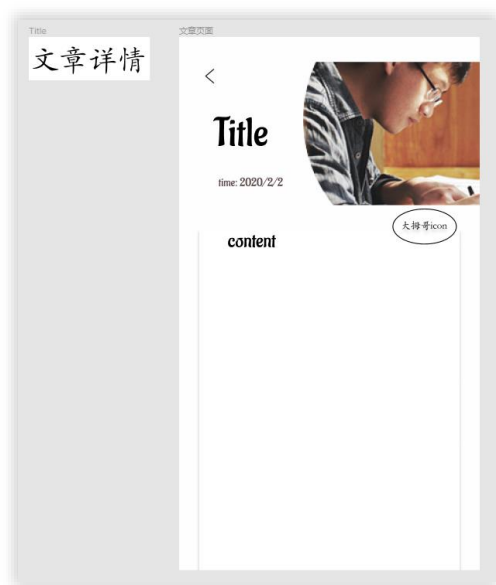


图 4.4-2 文章详情设计图

点击首页「开始拍照」的 button 后进入 3).

3) 拍照页面与整体性书写建议:

此页面为主要功能的入口页面，这是进入方式之一，另一种方式是从图库中直接选择图片。整体书写建议页面既是对于整体排版的结果反馈，也是选择单字的入口，从这里不仅可以看整体结果，也可以进入详细的单字纠正和评判页面。

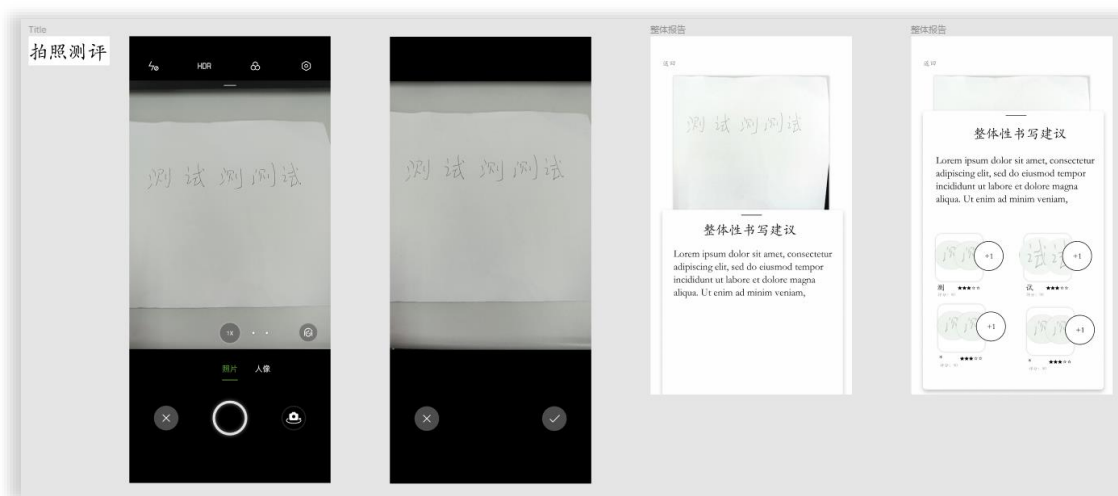


图 4.4-3 拍照评测设计图

4) 单字选择:

通过将用户选择的字高亮，从而方便用户选择想要查看的是这个字的哪一个字（对重复的情况容错），点击高亮框之后进入单字详情。



图 4.4-4 单字选择设计图

5) 单字详情:

此页面展示单字详情，包括检测字的骨架，标准的字的骨架蒙版，本字的详细参数（笔画数量、笔画长短、笔画偏移、整体结构）等，同时展示词字的评测历史记录。



图 4.4-5 单字详情设计图

6) 侧边栏与个人中心:

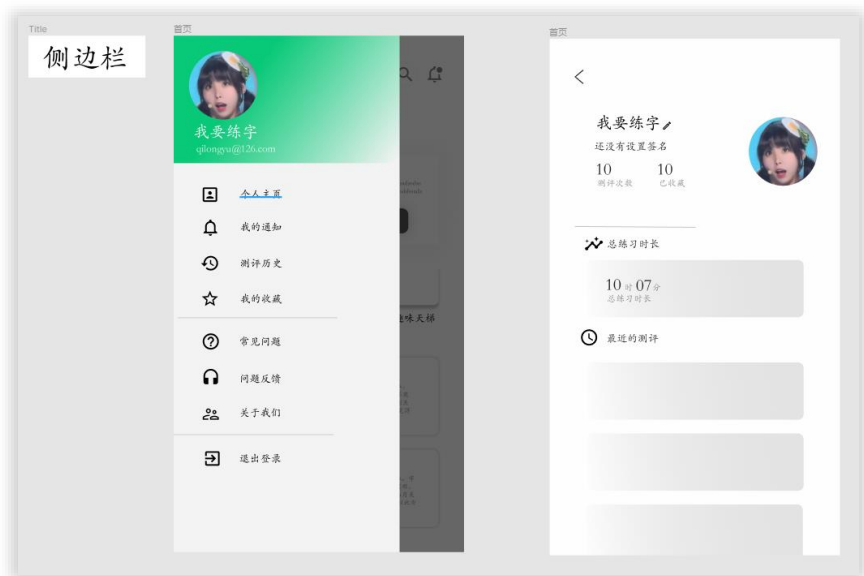


图 4.4-6 抽屉菜单设计图

7) 练习推荐:

生成两组推荐练习的字: 一组是推荐算法智能推荐, 另一组是通过分析该用户之前练习哪个笔画效果最差则刻意练习 (包括练习比较少的笔画, 增加权重)



图 4.4-7 练习推荐设计图

8) 深度追踪:

本页面通过两种方式追踪用户之前的评测记录。以字为单位, 通过记忆曲线, 帮助用户寻找之前评分较差的字, 通过重复练习来达到巩固和提升。另一种方式为按照日期和时间轴的方式来追踪用户的练习情况。

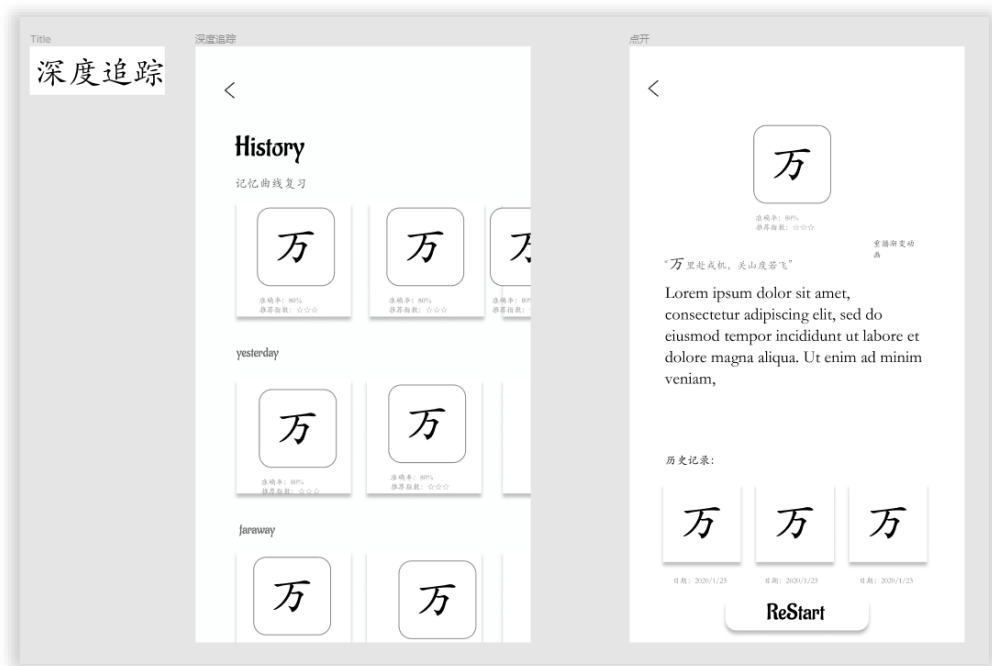


图 4.4-8 深度追踪设计图

9) 字体定制:

此功能为拓展功能，目的是通过利用用户之前的评测记录，利用风格迁移算法，迁移到目前用户想生成的文本上。



图 4.4-9 字体定制设计图

4.5 接口需求

4.5.1 硬件接口

无

4.5.2 软件接口

AIUnit 提供了文本框识别的业务，根据本项目在手写字识别任务上的需求，可以使用 AIUnit 提供的文本框识别业务来初步识别手写字的文本框，减少服务器运行压力，有效利用手机端计算资源。

4.6 其他需求

可维护性

系统的可维护性是衡量一个系统的可修复(恢复)性和可改进性的难易程度。

落笔云烟的可修复性主要体现在下面几个方面：

1. 数据库的可恢复性：我们将使用另一台服务器作为数据库的备份，并且使用了 cron 的任务调度，每天会定时将数据库文件数据备份到另一台服务器。
2. 完善的日志系统：Flask 轻量级框架可以实现完善的日志系统，落笔云烟利用了这一特性，会实时记录应用的运行日志，方便日后对于系统的维护。

可改进性

落笔云烟的可改进性体现在以下几个方面：

1. 技术的可改进性：落笔云烟的主要功能（图像分割、字体打分等）结合了人工智能技术，人工智能相关技术的进步对于落笔云烟效果的提升具有很大帮助。
2. 功能的可改进性：落笔云烟的相关功能可以更具针对性地为某些独特的字体爱好者进行私人定制，大规模训练某一单一字体，以保证提供给意图练习该字体的用户更加有针对性的指导。

5 概要设计

5.1 处理流程

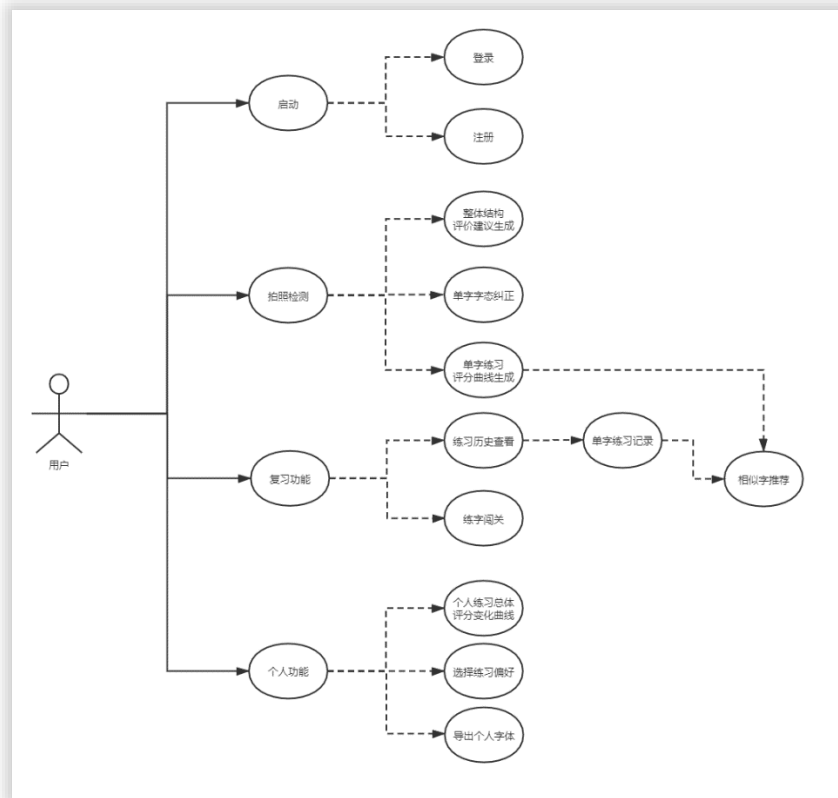


图 5.1-1 逻辑流程图

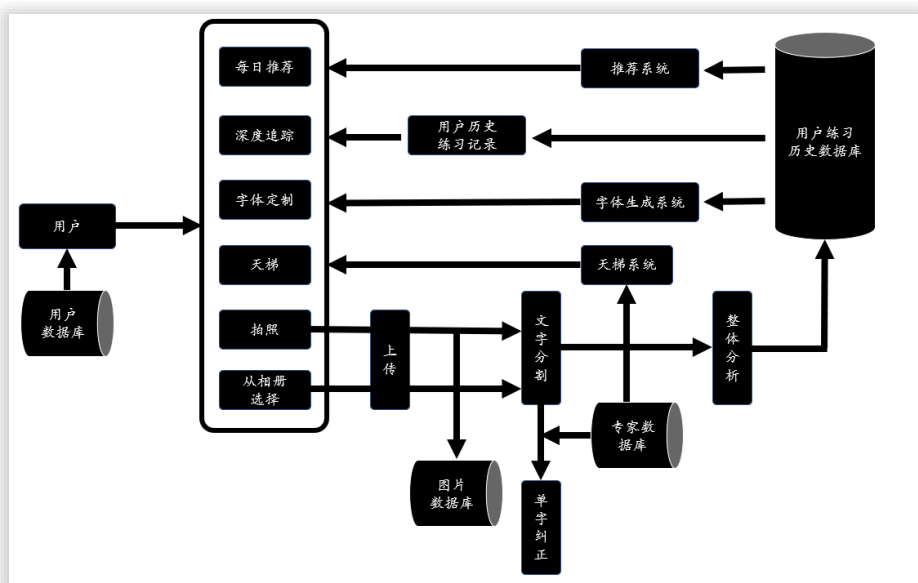


图 5.1-2 数据流程图

5.2 总体结构设计

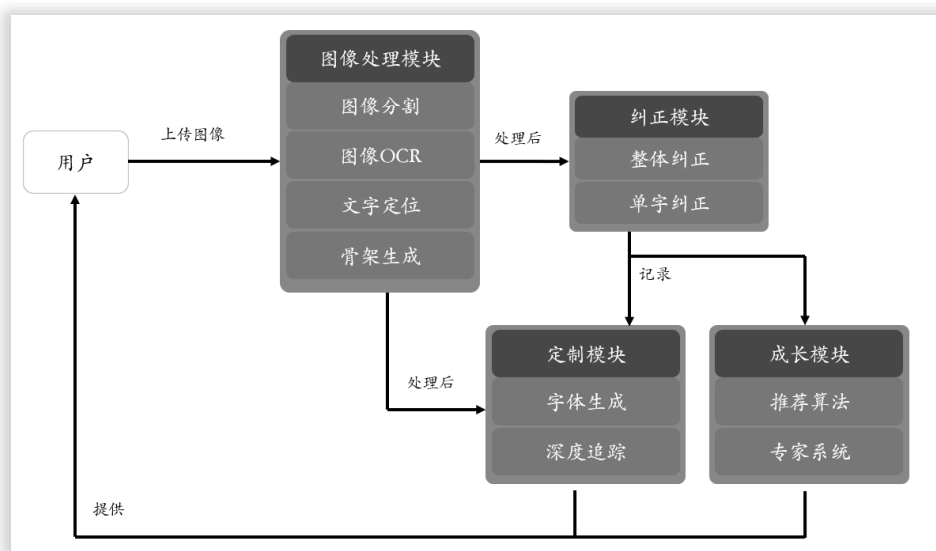


图 5.2-1 模块架构图

如图，落笔云烟的整体架构主要分为图像处理模块，纠正模块，定制模块以及成长模块四个部分。其中图像处理模块主要涉及图像处理技术，如图像分割、莫表检测、文字定位及 OCR 以及自研算法 OpenSkelenton，在整个架构中占据中枢位置；纠正模块主要包含整体纠正和单字纠正两个部分，主要涉及基于文字骨架的文本纠正算法；成长模块主要包含推荐算法以及专家系统；定制模块包含用户个人风格字体生成以及历史记录追踪两个部分，主要涉及风格迁移、特征分类等相关技术。

5.3 功能设计

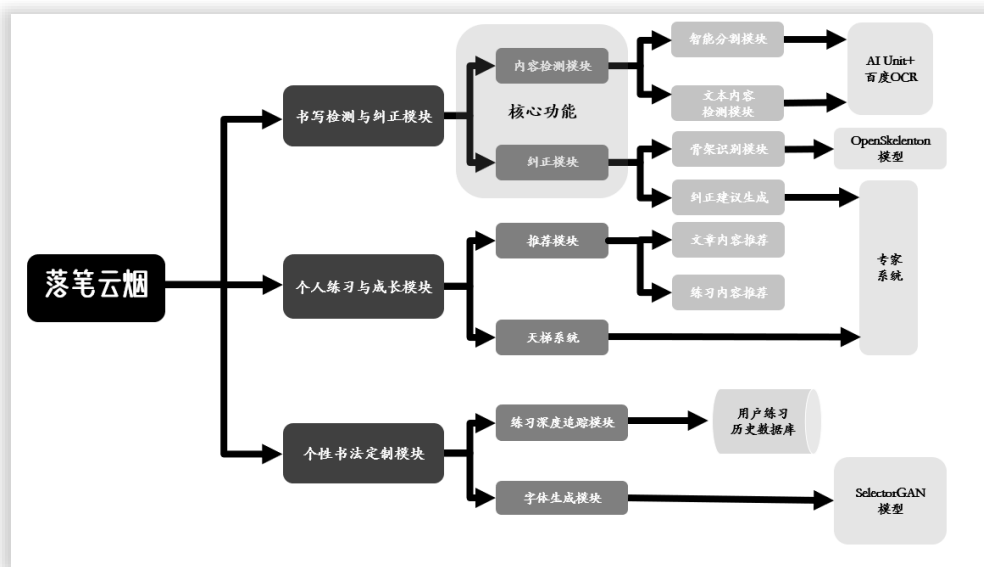


图 5.3-1 功能架构图

功能模块结构图如上，主要包含书写检测与纠正、个人练习与成长、个性书法定制三大模块。

1. 书写检测与纠正模块作为核心功能模块，主要围绕汉字骨架识别的 OpenSkeleton 算法对用户的书写内容进行分析，并依靠专家系统为其提出合理建议，改善用户的书写习惯，让书写更加协调美观；系统有机结合了 AIUnit 提供的端侧算力支持，算法可以更迅速地得到用户书写内容，极大提升用户使用体验。
2. 个人练习与成长模块。此模块主要基于协同过滤算法为用户提供个性化的学习和互动内容，并根据用户的练习情况为其匹配相应的天梯等级，为枯燥的练习过程增强一定的竞技性。
3. 个性书法定制模块。此模块主要包含以用户练习历史为基础，当用户的练习达到一定量之后，系统可以根据用户的书写数据进行分析，借助对抗神经网络模型 SelectorGAN 进行客制化字体生成，给用户提供阶段性的练习反馈，增强练习的趣味性；深度追踪模块通过追踪用户练习记录，对用户的易错点进行分析，并据此为用户准备练习和复习内容，从而起到对症下药、查缺补漏的作用。

5.4 用户界面设计

5.4.1 设计原则

落笔云烟团队一直以来都秉持一个信条,“设计绝对不是简单的拼合,排列甚至编辑;设计是通过阐明,简化、明确、修饰,使之庄严,有说服力,甚至带一点趣味性,来赋予其价值及意义”。因此,我们制作了一系列的界面设计原则,并始终坚持遵循这些原则进行用户界面的设计。

(1) 明确性

对任何界面而言,“明确”是首要的也是最重要的一点。人们必须能够辨别出它是什么,才能有效地使用你设计出来的界面。绝对不能出现让用户困惑的地方。明确的界面能够给使用者进一步操作的信心。

(2) 交互性

界面的存在是为了让人和我们的世界产生互动。它可以帮助人们厘清,明白,使用,展示相互之间的关系。优秀的界面、清晰的反馈不但能够让我们做事有效率,还能够激发、唤起和加强应用与这个世界的联系。

(3) 注意力

用户注意力是非常宝贵的,不要在应用的周围丢一些容易令人分心的东西。尊重用户的注意力,不仅仅会让用户感到高兴,设计也会收获好结果。如果在界面设计中,用户使用是首要目标的话,那么尊重用户的注意力是先决条件。要不惜一切代价保护它。

(4) 掌控性

人们会在自己能掌控的环境中感觉最舒心,最放松。草率的软件应用不但剥夺了这种舒适性,还会迫使人们面对毫无预期的互动,困惑的流程和意外的结果。为了提高应用的掌控性,需要通过定期的梳理系统状态,描述因果关系,并且在每一步操作都给出及时的提示和反馈,让用户感觉每一步操作都在他的掌控中。

(5) 流程自然

为用户精心设计交互的下一步操作。预期用户下一步的交互是怎样的,并且通过设计将其实现。就像日常谈话,要为深入交谈开一个好头。当用户已经完成要做的操作后,不要让他们不知所措的停留在那。提供自然而然的下一步,帮助他们完成操作。

5.4.2 用户界面实例



图 5.4-1 登录/注册界面

登录/注册界面

首先，打开应用用户可以看到以 logo 为主体的 Splash 动画，同时读取后台 token，如果未读取到 token 则进入登录/注册界面；如果读取到 token 则直接进入

到首页。

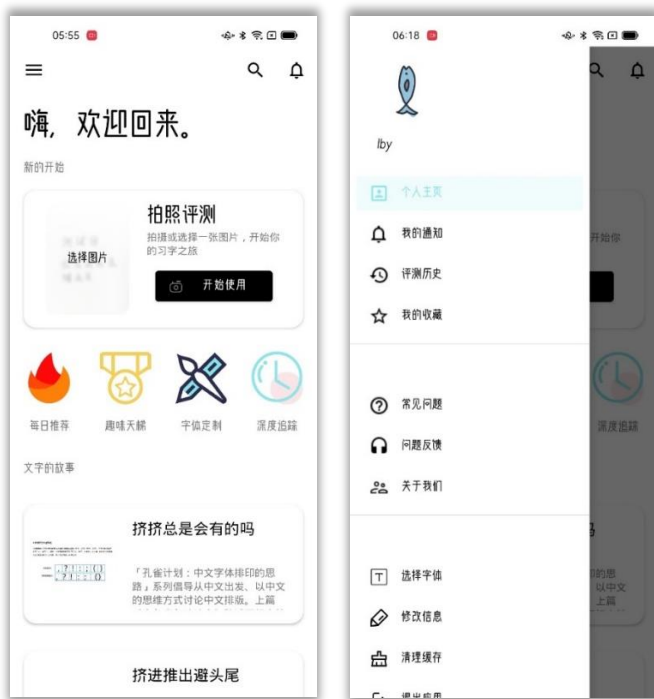


图 5.4-2 首页界面/抽屉菜单

首页

登录成功后进入首页，左上角可以拉起抽屉菜单，右上角可以查看通知提醒/直接搜索单字。在首页可以进入各大功能入口，点击“开始拍照”或者“选择图片”直接进入核心功能界面。

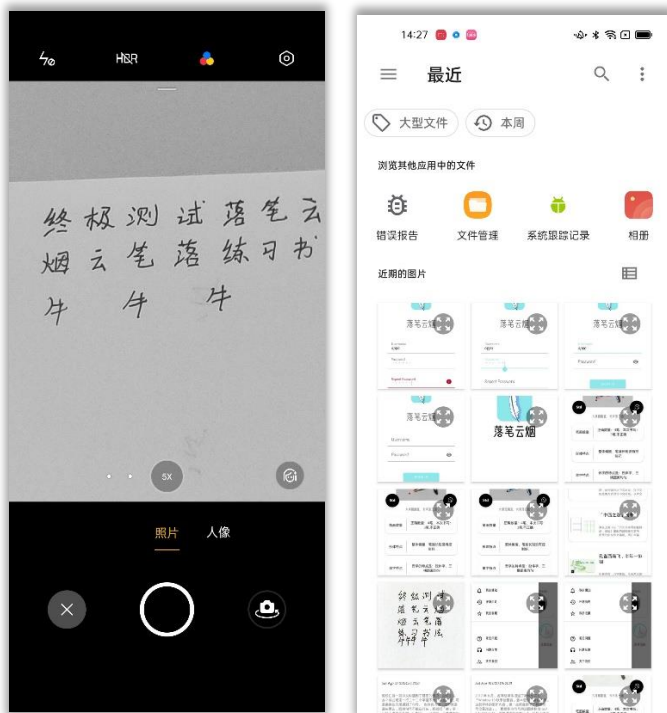


图 5.4-2 首页界面/抽屉菜单



图 5.4-3 整体评测界面

整体评测界面

在整体评测界面, 浮动卡片上显示了专家系统给出的整体书写建议。可以通过上滑浮动卡片显示所有识别到的文字, 文字下方可以查看当前字的整体评分, 点击文字可以对图片中出现的所有相同字进行定位, 并选择想进行纠正的文字进行后续的单字纠正工作。

单字纠正界面

在单字纠正界面, 可以查看该字的详细书写情况分析 & 系统给出的建议; 也可以通过点击文字两侧的浮动按钮进行手写字骨架和标准字骨架的显示 & 对比, 帮助用户更方便地定位到自己所写字的错误; 下方两个按钮可以切换到相同字的不同实例, 方便用户快速切换浏览并总结自己在该字上的整体错误。

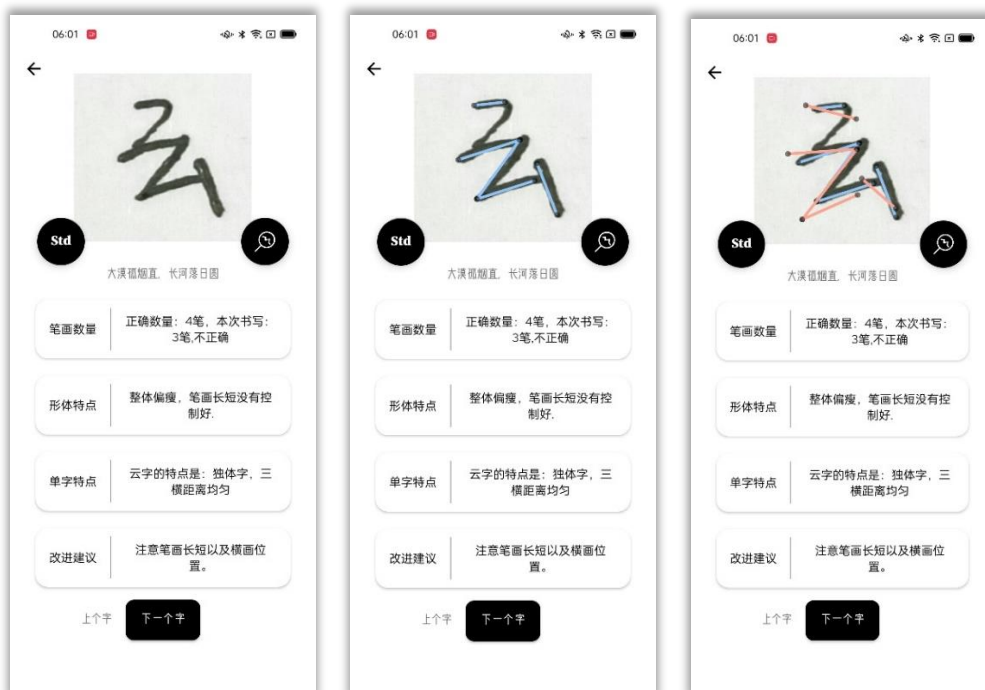


图 5.4-4 单字纠正界面

查看文章界面

在主页下部，我们提供了一些与汉字文化、排版等相关的精品文章供用户练习之余查看，在右下角用户可以为文章点赞，让更多的人看到这篇文章。



图 5.4-5 查看文章界面

5.5 数据结构设计

5.5.1 逻辑结构设计

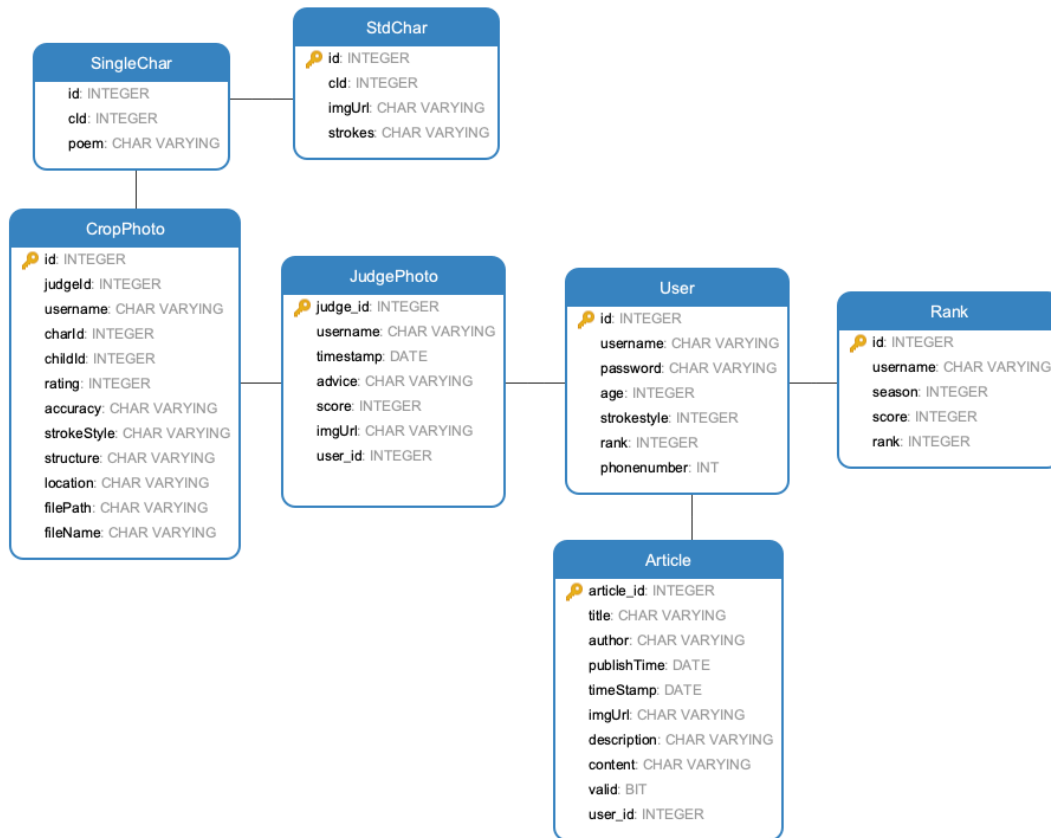


图 5.5.1-1 数据 ER 图

5.5.2 物理结构设计

数据库物理设计阶段的任务是根据具体计算机系统(DBMS 和硬件等)的特点,为给定的数据库模型确定合理的存储结构和存取方法。所谓的“合理”主要有两个含义:一个是要使设计出的物理数据库占用较少的存储空间,另一个对数据库的操作具有尽可能高的速度。在设计物理结构之前,落笔云烟首先对数据库的选择最初分析。对于关系型数据库来讲,其最大优点就是事务的一致性,这个特性,使得关系型数据库中可以适用于一切要求一致性比较高的系统中,比如:银行系统。但是在网页应用中,对这种一致性的要求不是那么的严格,允许有一定的时间间隔,所以关系型数据库这个特点不是那么的重要了。相反,关系型数据库为了维护一致性所付出的巨大代价就是读写性能比较差。对于本项目来讲,对并发读写能力要求极高且数据是海量的、增长的速度是难以预期的,因此采用非关系数据库 mongoDB 作为项目数据库。

落笔云烟在数据库的物理结构上主要做了以下的设计:首先,保证大部分数

据表满足 3NF 范式约束，降低数据之间的冗余；其次，在某些需要进行连接操作的数据表中，鉴于 NoSql 对于连接操作的不支持，项目中允许引入部分冗余键来保证查询的速度；最后，通过集群来实现负载均衡

5.6 接口设计

5.6.1 外部接口

软件接口：百度 AI 开放平台——手写文字识别

接口描述：对图片中的手写中文、手写数字进行检测和识别，针对不规则的手写字体进行专项优化，识别准确率可达 90% 以上。

请求说明：

HTTP 方法：POST

请求 URL： <https://aip.baidubce.com/rest/2.0/ocr/v1/handwriting>

请求参数：

表 5.6.1-1 手写文字识别接口请求参数表

参数	值
Access_token	通过 API Key 和 Secret key 获取 access_token
Content-Type	application/x-www-form-urlencoded
image	图像数据，base64 编码后进行 urlencode, 要求 base64 编码和 urlencode 后大小不超过 4M，最短边至少 15px，最长边最大 4096px，支持 jpg/jpeg/png/bmp 格式
recognize_granularity	是否定位单字符位置，big: 不定位单字符位置，默认值；small: 定位单字符位置
probability	是否返回识别结果中每一行的置信度，默认为 false，不返回置信度
detect_direction	是否检测图像朝向，默认不检测，即：false。朝向是指输入图像是正常方向、逆时针旋转 90/180/270 度。可选值包括：true: 检测朝向；false: 不检测朝向

● 返回说明：

返回参数：

表 5.6.1-2 手写文字识别接口返回参数表

字段	是否必选	类型	说明
log_id	是	uint64	唯一的 log id，用于问题定位
words_result_num	是	uint32	识别结果数，表示 words_result 的元素个数
words_result	是	array[]	定位和识别结果数组
location	是	object{}	位置数组（坐标 0 点为左上角）
left	是	uint32	表示定位位置的长方形左上顶点的水平坐标
top	是	uint32	表示定位位置的长方形左上顶点的垂直坐标
width	是	uint32	表示定位位置的长方形的宽度
height	是	uint32	表示定位位置的长方形的高度
words	是	string	识别结果字符串
chars	否	array[]	单字符结果， recognize_granularity=small 时存在
location	是	object{}	位置数组（坐标 0 点为左上角）
left	是	uint32	表示定位位置的长方形左上顶点的水平坐标
top	是	uint32	表示定位位置的长方形左上顶点的垂直坐标
width	是	uint32	表示定位位置的长方形的宽度
height	是	uint32	表示位置的长方形的高度
char	是	string	单字符识别结果
probability	否	float	当请求参数 probability=true 时返回该字段，表示识别结果中每一行的置信度值，包含： - average：行置信度平均值 - variance：行置信度方差 - min：行置信度最小值
direction	否	int32	图像方向，当

			detect_direction=true 时存在 -1:未定义, 0:正向, 1: 逆时针 90 度, 2:逆时针 180 度, 3:逆时针 270 度
--	--	--	--

5.6.2 内部接口

表 5.6.2-1 内部接口 URL/功能对照表

模块	接口 URL	功能描述
用户账户 设置	/api/u/register	用户注册
	/api/u/login	用户登录
	/api/u/set-password	用户修改密码
	/api/u/info	获取用户设置等数据
首页	/api/h/articles	获取首页展示文章数据
书写检测 与纠正	/api/f/main/ocr/preprocess	调用 AIUnit 的 Ocr 模块对图像进行预处理后将数据上传
	/api/f/main/single/char	获取单字的建议、评分、骨架数据以及标准字的图片、骨架数据
	/api/f/main/ocr/picture	上传用户拍摄/选取的照片, 获取图片中所有的文字数据及其位置信息
个人练习 与成长	/api/f/train/get/rank/list	获取用户天梯列表数据
	/api/f/train/update/rank	更新用户天梯状态
	/api/f/train/recommend	获取每日文字推荐数据
个性书法 定制	/api/f/assist/font/custom	获取用户定制字体
	/api/f/assist/history/char	按照单字获取用户练习历史记录
	/api/f/assist/history/all	获取用户所有练习历史

5.7 错误/异常处理设计

错误和异常的处理和反馈, 会一定程度上影响软件的调试效率、性能表现甚至用户的使用体验。因此, 落笔云烟在错误/异常处理方面进行了一定程度的规

范。

5.7.1 错误/异常输出信息

表 5.7.1-1 错误/异常输出信息表

错误类型	错误位置	输出信息
网络错误	图片上传	图片上传失败
	图片加载	图片下载失败
	登录/注册	登录/注册异常, 请检查是否已接入互联网
权限错误	登录失效	登录已过期, 请重新登录
	请求接口数据	请先进行登录!
数据错误	图片上传	无法检测到文字, 请重新拍摄
用户输入错误	注册	两次密码输入不一致/密码强度弱
请求错误接口	请求接口数据	请求接口错误

除在客户端和后台输出错误信息外, 对 Nginx, flask 与客户端后台均配置有错误日志系统; nginx 日志用于收集错误请求, flask 用于收集模型、算法以及数据库错误信息, 方便调试时进行查看。同时通知相关线程释放资源, 避免资源浪费; 同时返回给接口调用者错误类型信息使其变更操作方式或通知用户, 从而保证用户的良好交互性。

5.7.2 错误/异常处理对策

表 5.7.2-1 错误/异常处理对策表

错误类型	错误位置	处理对策
网络错误	图片上传	客户端提示用户接入网络, 并抛出异常, 集中进行异常捕获, 并重启当前活动
	图片加载	
	登录/注册	
权限错误	登录失效	重新获取 token
	请求接口数据	

数据错误	图片上传	模型内部抛出异常，并捕获异常进行处理
	数据格式错误	服务器抛出异常，并返回 500 错误
用户输入错误	注册	在客户端的数据模型中进行验证，并对用户进行即时提醒
请求错误接口	请求接口数据	服务器拒绝访问，返回 404 错误

除上述错误外，当进行其他敏感操作时，可能会引发预料之外的错误，此时处理不当可能会对数据库和用户数据造成较大的影响。因此，需要定期进行冗余备份，并提前设置好后台服务器宕机重启后的实例恢复、数据恢复等操作。

5.8 系统配置策略

落笔云烟利用 Flask+Celery+Redis+MongoDB 组合作为业务处理逻辑，采用 Nginx+Gunicorn 进行部署。

- 采用 Flask 的原因

1. Flask 框架为轻量级 Web 应用框架，其有着极强的扩展性，对比 Django 等框架，Flask 拥有更高的灵活性适合自己设计代码框架。
2. 由于本项目需要使用多种自实现深度学习模型，使用 Python 应用框架便于加载模型

- 使用 Redis 的优势

1. 对于纯静态数据，直接使用 URL 进行缓存，在之后处理中可直接返回数据，无需访问磁盘。
2. 采用 Redis 的另一个目的便是作为 Celery 任务队列的中间件。

5.9 系统部署方案

本项目基于 Flask 框架开发，采取 nginx+wsgi+flask 技术路径进行部署。由于 Flask 是基于同步进程的后端框架，处理请求时以单进程方式进行响应，当短时间内产生大量并发请求时，Flask 服务会出现阻塞。因此，原生 Flask 框架的 web server 性能尚不能满足项目要求。为解决此问题，落笔云烟利用 nginx + gunicorn + flask 方式进行部署，采用更高性能的 wsgi server 代替 Flask 原生 server，以取得更优秀的并发性能。

其他方面,项目使用 Nginx 将静态数据请求转发至静态文件服务器进行处理;同时, Gunicorn 依赖 Nginx 的代理行为,与 Nginx 进行功能剥离,仅处理动态请求,并在处理完毕返回数据给 Nginx,再由其发送回客户端。最后,在架构层面,项目还增加了守护进程 supervisor 作为监控服务进程的解决方案。

5.10 其他相关技术与方案

1. 汉字笔画数据集标注工具: (自行开发)

业务需求: 在模型中,汉字被解构为笔画信息,每个笔画又被结构为了关键点。为了实现“将汉字解构成关键点,组成笔画,从而进行汉字骨架识别”的算法要求,在数据集方面,我们需要录入关键点来表示一个汉字,即汉字→关键点的数据集。

- 1) **解决思路:** 由于目前可访问的网络资源中没有符合要求的数据集,于是团队目前的解决方案是采用 CASIA 中科大手写数据集,以及从网站爬取的标准 ttf 字体文件作为基础训练集图片,在此基础上进行标注。为此我们开发了一款专门用以标注汉字结构信息的 web 应用,通过爬取所有汉字的笔画信息,使程序自动读取汉字笔顺,结合人工标注,以半自动的方式运行。现我们已将应用开源,录入效率约为 **5.32 字/分**。
- 2) **技术方案:** react.js/react hook/nodejs/express
- 3) **效果展示:**



图 5.10-1 数据集标注工具

2. 目标检测数据集制作

- 1) **业务需求:** 训练从纸张上识别出人类手写的汉字。数据集需求: 写满字的

A4 纸 + 每个字的包围盒坐标值。

- 2) **解决思路:** 考虑到 CAISA 的数据集为已经被分割出来的汉字, 我们考虑将其拼回一张 A4 纸上, 巧妙地反向生成我们所需要的数据集。
- 3) **技术方案:** 使用 node.js , ImageMagic Display 与 npm gm, 读取原始数据集, 拼接回 A4 纸内, 并按照格式详细记录拼接信息用于算法训练。
- 4) **结果展示:**

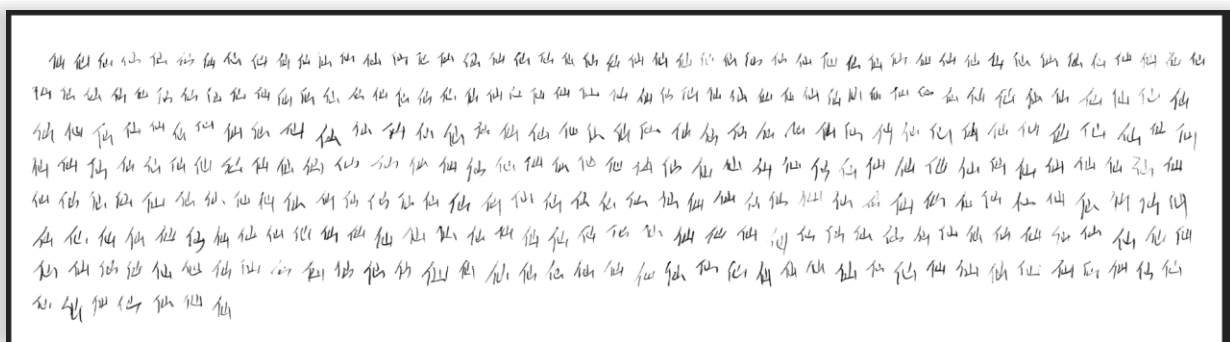


图 5.10-2 目标检测数据集生成

6 数据库设计

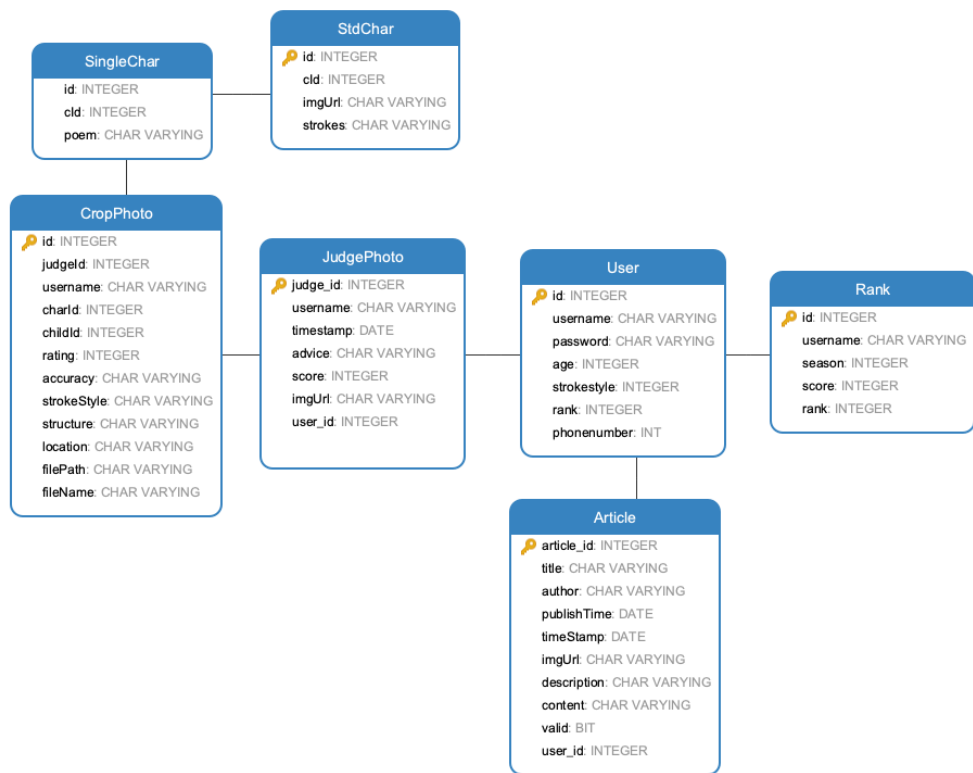


图 6-1 数据库 E-R 图

7 手机端侧部署设计

7.1 手机环境需求

要求采用 oppo 系列手机，系统要求 ColorOS11+；安装有 AIUnitServer 支持框架并开启手机关联启动项。

7.2 调用 AI Unit 能力接口设计（可选）

项目采用 AI Unit 的文本框识别 OCR 业务，对用户选择即将上传的图片进行文本与背景的粗分割（由于 AI Unit 无法进行参数调整，因此无法控制模型粒度，在采用目前默认模型参数的环境下，仅能做到词组分割），获得 AI Unit 结果的 BoundingBox 数据，通过上传接口提交到后端服务器，方便服务器在已分割完毕的词组数据的基础上，对单个文字进行后续的目标检测及分割处理，从而极大减轻后台模型的目标检测负担。

7.3 AI Boost 端侧模型加速设计（可选）

8 详细设计

8.1 书写内容分割模块

8.1.1 功能描述

给定一张用户书写的图片，识别出图片中的所有手写的汉字的种类和边框，将检测出的汉字信息与边界框交给其它模块处理。

8.1.2 性能描述

一张图片的目标检测需要在 1s 内完成，当服务器请求过多时，可以使用 GPU 的并行计算性，同时完成多张图片的检测，用户端的响应时间应在 1s 以内。

8.1.3 输入

表 8.1-1 书写内容分割模块输入示例

序号	内容	类型
1	用户书写的图片	图片

8.1.4 输出

表 8.1-2 书写内容分割模块输出示例

序号	内容	类型
1	用户上传的原始图片	图片
2	识别出的汉字	字符串
3	识别出的汉字的边界框	数组

8.1.5 程序逻辑

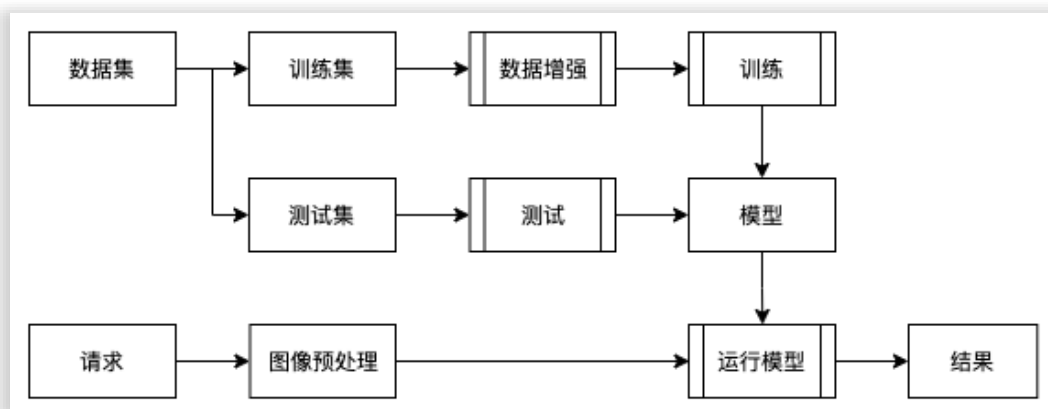


图 8.1 书写内容分割模块功能逻辑

8.1.6 限制条件

书写汉字不能有重叠，否则影响识别率，拍摄图片需保证充足光照，否则影响二值化效果。

8.2 书写字形纠正功能模块

8.2.1 功能描述

改模块分为汉字骨架识别与汉字字形纠正两部分，汉字骨架识别通过识别出汉字的笔画信息提供给汉字字形纠正部分作为基础信息。汉字纠正部分根据识别出的汉字骨架，通过计算用户写的汉字与模版汉字的各个笔画的区别，判断各个笔画的书写情况。之后通过数据库中预先建模好的汉字的结构信息，匹配该汉字存在的书写缺陷及指导建议反馈给用户。

8.2.2 性能描述

手写字识别的运行时间在 3s 之内，字形纠正的运行时间在 1s 之内，总的响应时间在 5s 之内。

8.2.3 输入

表 8.2-1 书写字形纠正功能模块输入示例

序号	内容	类型
1	单字的图片	图片
2	识别出的汉字 ID	Int

3	模版汉字 ID	Int
---	---------	-----

8.2.4 输出

表 8.2-2 书写字形纠正功能模块输出示例

序号	内容	类型
1	输入汉字的骨架信息	数组
2	模版汉字骨架信息	数组
3	输入汉字调整后的图片	图片
4	模版汉字调整后的图片	图片
5	输入汉字与模版汉字之间的笔画的不同点	字符串
6	输入汉字存在的缺陷问题和指导信息	字符串

8.2.5 程序逻辑

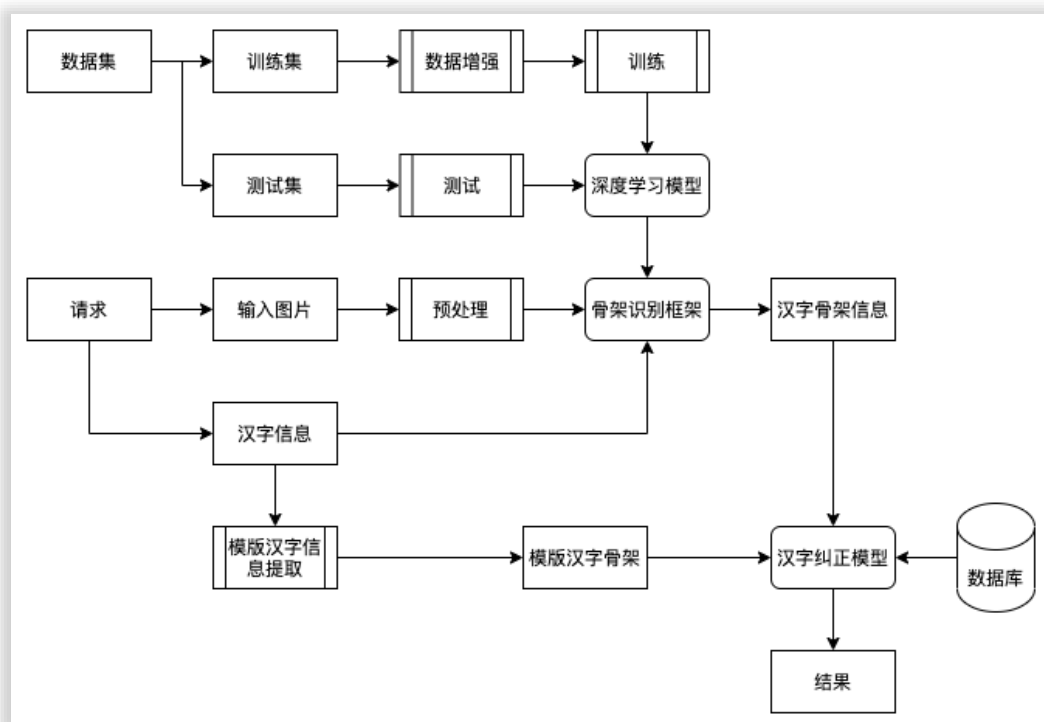


图 8.2 书写字形纠正功能模块功能逻辑

8.2.6 限制条件

待识别的汉字需要存在于训练汉字骨架的模型的数据集中。数据库中需要包

含待纠正的汉字的相关信息。

8.3 兴趣培养与成长功能模块

8.3.1 功能描述

在传统理解中,练字本身是一件枯燥且耗时得事,且由于见效慢,缺乏正反馈,导致许多人难以长期坚持。因此,为了缓解这一现状,落笔云烟借用现如今常用于各类游戏中的天体系统,解决练字过程中缺乏正反馈得难题。

在落笔云烟应用中,规定每3个月为一个赛季。每赛季中,为每个用户设定初始分值,同时每周进行百人练字大赛,通过对规定文章临摹进行打分,进而为参加比赛的用户进行排名,设定段位值。与此同时,为了令用户在比赛中有着更好的表现,落笔云烟会依据用户历史记录为用户提供对应的书写练习。

8.3.2 性能描述

- 天梯系统:
无性能需求。
- 个性推荐:
整体推荐算法计算时间<1500ms,内存占用<100MB。

8.3.3 输入

- 天梯系统:
对每个用户输入对应的练习稿。
- 个性推荐:
无输入。

8.3.4 输出

- 天梯系统:
输出该用户对应分值以及该轮比赛其所获得加分。
- 个性推荐:
输出该用户的推荐练习。

8.3.5 程序逻辑

- 天梯系统:

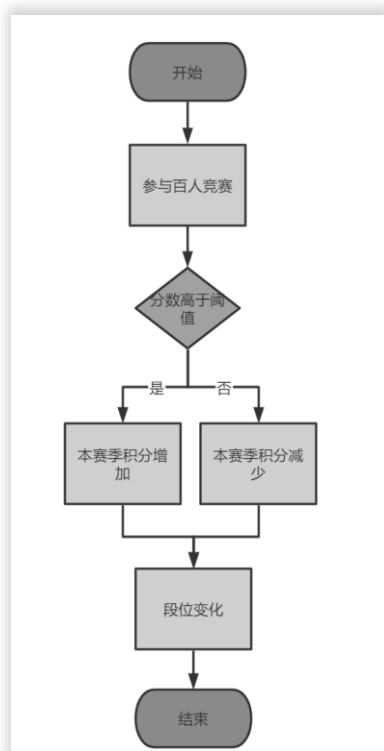


图 8.3-1 天梯系统功能逻辑图

- 个性推荐:

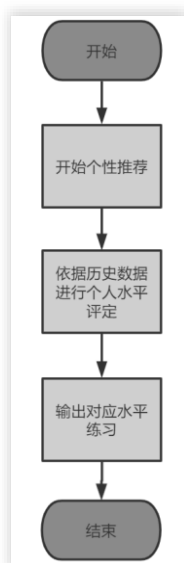


图 8.3-2 个性推荐功能逻辑图

8.3.6 限制条件

无

8.4 个性书法定制功能模块

8.4.1 功能描述

对于每个练字爱好者，谁不想有一套自己的字体集呢？落笔云烟将为这些用户提供这一定值服务。本应用将依托于用户练习记录，利用深度追踪学习技术，为用户提供个人字体生成服务。

8.4.2 性能描述

鉴于个性字体生成服务需要基于用户的字体风格以及大量练习记录，计算量较大，因此即使生成无法保证运行时间，因此落笔云烟将在用户历史练习量超过阈值后，为其数据进行自动分析，若用户之后发起生成请求，即可保证即时产生结果。请求等待时间<1000ms,请求内存消耗<100MB。

8.4.3 输入

输入与该用户相关的字体风格以及历史练习记录。

8.4.4 输出

基于该用户字体风格的字体集。

8.4.5 程序逻辑

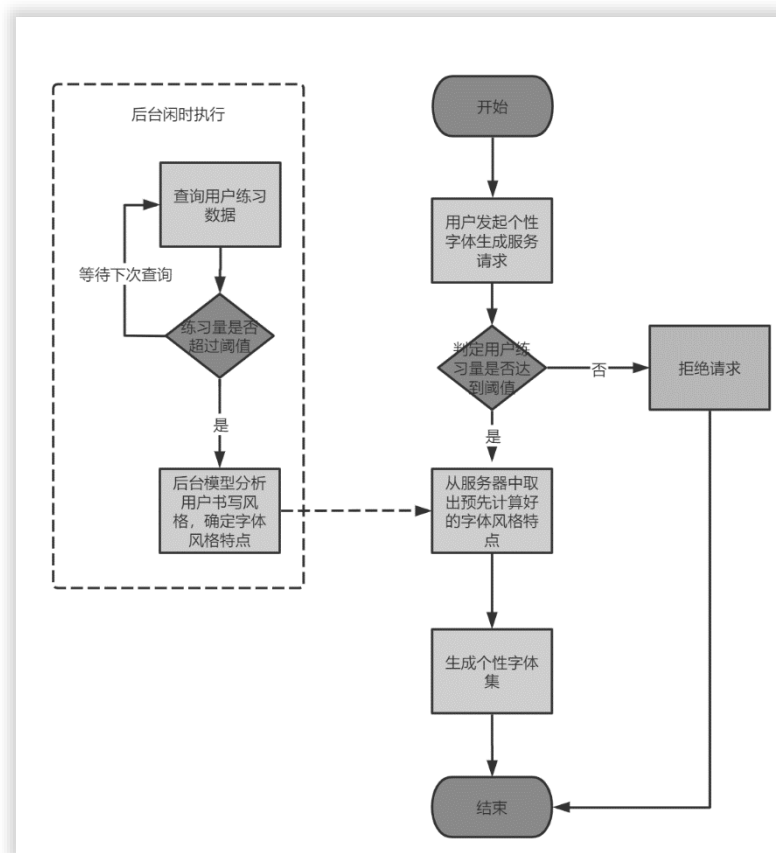


图 8.4-1 个性书法定制功能逻辑图

8.4.6 限制条件

基于前文的介绍可以看出,本功能最大的限制条件便是用户的训练条目数量。只有保证了一定数量的训练,模型才可以更加准确的为用户提供该服务。

8.5 **功能模块