

基本资料

姓 名：高甜甜

性 别：女

年 龄：32

学 历：本科

手 机：15001280512

邮 箱：1298021806@qq.com

毕业院校：河北科技师范学院

专 业：计算机科学与技术

职业意向

期望行业：算法工程师

期望地点：北京

个人技能

- 1、编程与工具：熟练掌握 Linux 环境及 Shell 命令，熟练使用 Python、Java；熟悉 Docker、Nginx 等工具。
- 2、机器学习：熟练机器学习算法，包括线性回归、逻辑回归、XGBoost、SVM、K-means 聚类等。
- 3、深度学习：掌握深度学习模型，如 ResNet、YOLO、PaddleOCR、Qwen 等。
- 4、深度学习框架：熟悉 TensorFlow、PyTorch、PaddlePaddle 等框架。
- 5、大模型应用：熟悉 LLM 应用开发技术，包括提示工程、RAG 架构、LangChain、Dify。
- 6、大模型微调与部署：熟悉 Ollama、vLLM、LMDeploy 及 LoRA 微调技术。
- 7、工具与平台：熟悉 Faiss、Selenium、Airtest 等工具。
- 8、数据库：熟悉 MySQL、Oracle、Redis。

工作经历

2022.10-2025.5

随行付

算法研发

工作地点：北京 石景山

职责业绩：

1、智能审单系统，包括同图检索、OCR 识别、多模态模型微调等。

2、内部提效，包括测试提效和研发人员提效。

3、人脸比对模型训练及优化。

4、其他模型探索及业务方案设计、对比，需求沟通等工作。

2017.08-2022.10

中能华云（北京）能源科技有限公司&北京赛思天成科技有限责任公司

图像算法研发+Java 开发

工作地点：北京 昌平

职责业绩：

1、危险行为检测、人脸识别、车牌识别模型训练及优化。

2、站级订货系统、云化 BOS、POS 的前后端开发，后端框架主要用到了 Springboot、SpringMVC、mybatis、springcloud，前端框架主要用到了 jQuery、thymeleaf、bootstrap、vue 等，权限控制用到了 shiro，数据库用到了 oracle、MySQL、sqlite。

2015.06-2017.07

中科弘睿科技有限公司&纵横聚牛科技有限公司&秦皇岛商之翼网络科技有限公司

Java 开发

工作地点：北京、河北

职责业绩： 1、 开发医院系统，主要负责系统的前后端开发、插件编写、功能优化，主要使用了 Spring、mybatis、Spring MVC 作为框架，使用 js、jQuery、HTML、css 编写前端页面及插件，数据库用到了 oracle 和 MongoDB，利用 Echarts、IReport 开发报表。

2、 使用 HTML、css、js、jQuery 编写公司网站页面。

3、 在 ECShop 模板的基础上根据公司需求，开发新的模板，主要使用 PHP、js、jQuery、HTML、CSS。

## 项目经历

### 基于深度学习和大模型的智能审单系统

项目职责： 数据处理、模型训练以及方案选择，提升审核效率

项目描述： 项目背景：

公司需要对商户的资质、法人信息等进行审核，为了提高审核人员的效率，加入了 OCR 识别和 mcc 智能推荐功能，为了减少合规风险，加入了同图检索功能，防止多个商户使用相同的门头和内景照。

其中 mcc 推荐业务分为两种情况，一类商户需要先根据传入的门头内景照，判断是否在给出的经营范围内，之后再根据图片进行 mcc 推荐；三类商户直接根据商户名称和经营范围给出 mcc 推荐。

工作内容：

#### 1、 经营范围匹配：

根据业务场景，最终引入多模态方式，对比多个模型，选择使用 doubao1.5-vision-pro-32k 模型，准确率提升至 80%以上。后续尝试使用 lora 微调 qwen2.5-vl-7B 模型，准确率约 85%，为后续优化奠定基础。

#### 2、 三类商户 mcc 推荐：

采用 RAG 架构，使用 moka-ai\_m3e-base 模型实现文本到向量转换。对比 faiss 和混合检索方式，最终采用 jieba+混合检索+大模型的方式，准确率提升至 70%以上。后续又尝试了使用大模型进行分词，最终效果和 jieba 分词差不多，但是速度会慢一些，没有采用。

#### 3、 同图检索：

对数据进行处理，最终构建了 2 万多组图片的训练集。使用 resnet50 训练模型，将图片转换为 128 维向量，减少内存占用。使用 faiss 进行图片检索，同时存储到 Redis，支持百万级图片库，每天凌晨更新数据，准确率约 99%。

#### 4、 银行卡 OCR 识别：

之前离职同事项目，准确率已有约 90%，通过优化检测模型和过滤脏数据，整体准确率提升约 5%。

#### 5、 对公开户证明 OCR 识别：

使用 YOLOv7 结合 paddleOCR 进行训练，此场景有几十种样式，经过几轮优化，最终准确

率达到约 85%，后续生产上还有新样式加入，目前生产准确率平均 75%左右。

## 6、授权函 OCR 识别：

使用 YOLOv7 结合 paddleOCR，但因为每张图有 19 项内容需要识别，且为手写文字，优化两轮后，整张准确率依然很低，项目暂停。

### 测试提效与研发提效

项目职责： 测试提效与研发提效

项目描述： 项目背景：

公司希望提高测试人员和研发人员的工作效率。

工作内容：

#### 1、测试提效：

开发造数工具，支持测试人员完成注册、认证等业务流程；推广 UI 自动化工具，以及 selecniium 控件封装、airtest 工具推广，降低测试人员代码编写难度；大模型使用推广。

#### 2、研发提效：

选择并部署适合代码生成的模型，对比 codeqwen、codellama、deepseek-coder 等模型，最终选用 deepseek-coder；优化部署方式，从 ollama 到 Docker，最终采用 vllm，使用 nginx 进行转发，提升并发性能。

### 基于深度学习的人脸对比

项目职责： 人脸对比模型训练与优化

项目描述： 项目背景：

为了提高 app 实名认证的安全性，需要增加人脸比对功能，判断进行人脸识别的用户是否和上传的身份证为同一人，准确率要达到 97%。

工作内容：

1、对 2 万组人脸照片进行处理，通过数据增强提升模型泛化能力。

2、尝试使用多种模型（deepface、arcface 等）进行测试和训练，最终使用 YOLOv8-face 进行人脸检测，结合修改后的 resnet50 进行训练，最终准确率提升至 97%。

### 基于深度学习的智慧加油站项目&人脸识别

项目职责： 模型训练及选择

项目描述： 项目背景：

智慧加油站项目：为了防止有人在加油站抽烟、打电话，因此开发危险行为检测功能，减少安全事故；为了使加油以及支付更加方便，因此训练车牌识别模型，通过车牌在后台查询出会员信息。

人脸识别：疫情期间，为了记录人员进出，训练人脸识别模型。

工作内容：

- 1、危险行为检测：使用 YOLOv5 进行训练，准确率约 95%。
- 2、车牌识别：使用 MobileNetV3+PaddleOCR 模型进行训练，经测试准确率可达到 96%左右。
- 3、人脸识别：使用开源的 MTCNN 结合 FaceNet 模型，经测试，准确率可达到约 96%。

## 自我评价

工作认真、负责，有问题积极沟通，有良好的团队合作精神。