

张腾

 johnson-magic ↵

 158-2135-6114

 Teng_Zhang ↵

教育经历

东华大学（211工程） 学士

GPA 排名 9/116

2011.09 – 2015.06

信息科学与技术学院, 自动化系

上海交通大学 硕士

导师 朱兰娟

2015.09 – 2018.07

电子信息与电气工程学院, 自动化系

- 会议论文：基于迁移学习的细粒度车型分类算法研究 中国过程控制会议 (cpcc) oral
- 毕业论文：基于卷积神经网络的细粒度车型识别系统研究与实现

工作经历

海康威视研究院（上海）

机器视觉算法工程师

2018.07 – 2020.11

智能算法部

- 工作内容：车辆特征信息结构化提取相关算法研发工作
- 所获荣誉：2019 年度公司创新奖《AI 自研芯片项目》；集体荣誉

上海合合信息科技股份有限公司（代表产品：扫描全能王，启信宝，名片全能王）

机器视觉算法工程师

2021.01 – 2024.08

AIM/机器视觉组

- 工作内容：(1) “扫描全能王” App 版面还原功能相关算法研发工作
(2) “蜜蜂家校” 小程序智慧教育相关算法研发工作
- 所获荣誉：(1) 2021 年季度最佳技术分享官：《从 focal loss 到 gfocal loss》；个人荣誉
(2) 2023 年度优秀团队《AI 批改服务小组》；集体荣誉

上海交通大学

科研助理

2024.09 – 至今

自动化系/VisionXLab ↵

- 工作内容：(1) 独立负责 0 到 1 搭建 mini 版 ocr 引擎 (与 卢长胜博士 合作)
(2) 负责 point-supervised 有向目标检测方向研究 (与 杨学教授 合作)
- 学术产出：(1) Point2RBox-v3 ICLR2026 已接收，共同第一作者（排名第一）

项目列表

海康威视研究院期间

2018.07 – 2020.11

- 项目名称：车辆特征信息结构化

项目时间：2018.07 – 2020.11

工作内容细分：

- (1) 负责：优化“一盔一带”等智慧交通应用的模型指标

背景：检测司乘人员是否佩戴头盔和安全带是智慧交通的重要功能
挑战：安全带具有遮挡、与座椅背景色相近等挑战，对特征的提取能力要求较高
行动：(a) 升级 yolov2 至 yolov3 方案
 (b) 采用 iou loss 提升检测框的准确度
 (c) 采用 sk、se 等视觉 attention 增强视觉特征提取能力
结果：海康“一盔一带”功能始终位于业界第一，超越大华、华为等友商
(2) 负责：基于公司自研 AI 芯片完成所在小组相关业务模型的迁移、适配
背景：由于华为海思芯片短缺，公司业务要向其他 AI 芯片迁移
任务：与总部（杭州）芯片组相关同事协调配合，完成组内业务的迁移、适配
行动：(a) 基于 C++ 编程语言和 Caffe 推理接口完成业务库的原型开发
 (b) 根据自研芯片的限制，对已有的业务模型结构进行适配、调整
结果：已有产品技术方案从华为海思芯片迁移至海康自研芯片，该项目获得 2019 年度公司创新奖
(3) 参与：研发统一部署框架并在上海研发中心进行推广
背景：to B/G 业务存在较多的定制化需求，但通常可抽象为检测 + 分类的方案框架
任务：从杭州总部引进统一部署框架并在上海研发中心推广应用
行动：(a) 从杭州总部引进并进行宣贯和推广
 (b) 根据反馈，在已支持华为 hisi19、59 的基础上，补充开发 ma 芯片部署框架
结果：该框架实现了检测、分类等智能应用在 hisi19、59 和 ma 等 ai 芯片上无感、高效部署，组内 3 人以上应用该框架发布超过共计 10 次以上 ai 应用，该方案也成为其他组发布小型 AI 应用的主流框架

合合信息期间

2021.01 – 2024.08

项目名称：“扫描全能王”APP 版面还原项目

项目时间：2021.01 – 2022.09

工作内容细分：

(1) 负责：梳理并维护通用目标检测 pipeline

背景：合合内部自研了一个类 yolov6 的检测库 detection-pytorch，但实现相对粗糙，未经过严格的指标验证

任务：对齐主流检测算法的论文指标，寻找 detection-pytorch 中各个模块可能存在的实施缺陷

行动：按单元模块，划分为数据预处理、backbone、match-sample-loss、coco 评测和超参配置 5 部分，结合 unit testing 的思想一一对齐

结果：寻找到所有自研检测库的可能缺陷；可通过配置，复现 atss、gflv1、gflv2 和 auto-assign 论文指标

(2) 负责：优化版面还原子服务---版面元素检测模型的指标

背景：版面元素检测是扫描全能王（CS）核心功能版面还原的子模块

挑战：经统计，C 端业务场景，用户上传的文档照片中有 2.2%~7.7% 的比例为非 0 度拍摄，角度因素会导致检测模型出现“框不够准”和“类别混淆”的问题

行动：(a) 价值分析：将训练集和测试集通过脚本离线转正，训练模型测试指标，获得理论收益上限

 (b) 快速可行性分析：单独训练拍摄角度方向判断模型，评估角度预测准确率

 (c) 方案实现：采用 stn 模块链接“角度判断-旋转-检测”，实现 end-to-end 方案

 (d) 部署落地：解决 pt 模型导出时的 if-else TorchScript 问题

结果：(a) 版面元素检测自测集上获得指标提升：中文场景 1.1mAp@0.75:0.95、英文场景 1.6mAp@0.75:0.95、外文场景 0.6mAp@0.75:0.95

 (b) 将该方案推广至智慧教育产品线中的题目信息检测服务：题目类别指标提升 6.2prf@0.75、作答区域类别提升 1.3prf@0.75，最终的线上业务指标作业批改率提升 0.9 个百分点

 (c) 截至到 24 年 8 月，两条业务线的线上模型仍旧采用该方案（价值贡献寿命达到 2 年）

项目名称：“蜜蜂家校”小程序 AI 批改项目

项目时间: 2022.09 – 2024.08

工作内容细分:

(1) 负责: 维护、更新教育 OCR 服务

背景: 教育 OCR 服务包含文本检测、阅读方向判断、文本识别和公式识别等子模块, 是 AI 批改的支撑服务之一

挑战: (a) 批改引擎是众多服务和复杂业务逻辑的糅杂, 较难做到子模块指标的提升与系统级业务指标正相关

(b) 批改引擎 AI 子服务众多, 是合合信息推理显卡占用最多的产品线之一, 成本较高; 且批改高峰时段, 推理阻塞会导致批改失效率增加

(c) AI 批改业务逻辑在依赖 OCR 识别结果的同时, 强依赖字符点位的输出

(d) 教育场景手写字体、小学生宽松字体识别带来较大的挑战

行动: 针对挑战 (a), 设计“完美伪子服务”集成至批改引擎, 获得对应子服务在 100% 准确的情况下, 业务指标的理论上限, 从而获得系统优化的方向和子服务优化的优先级

针对挑战 (b), 选型 faster-transformer, 将其与公式识别模型 satrn 适配, 替换 libtorch 的推理 backend 集成至公司级别的 ocr-engine, 实现服务的加速

针对挑战 (c), 在 nvidia 官方 beam-search-cuda 不输出 timesteps 的情况下, 采用 forced-align 方法反推 CTC 的 timesteps, 实现字符点位的预估

针对挑战 (d), 建立数据飞轮流程, 并借鉴 nlp 中的 tokenizer 思想设计类 radical-based 的字符编码方式

结果: (a) 建立起了科学的“系统--子模块”优化方向确定机制, 并将该机制宣贯、推广至扫面全能王中的版面还原业务线

(b) nvidia A10 芯片上, 线上 ocr 服务的推理速度提升 40%, 主导推动实现该功能集成至公司级别的构建平台中, 目前已被 2 个二级研发部门 (AIM、ACG)、三个重要的产品线所使用 (AI 批改、图片转 word、pdf2markdown)

(c) 顺利获得字符点位的正常输出, 支撑业务侧“留痕”批改功能

(d) 通过数据飞轮, 填空题的批改率提升 4.4 个百分点 (单季度提升量); 通过新的编码方式, 小学生手写文本的 ctc “重影”问题得到充分的解决

(2) 负责: 探索并落地 LLM 大语言模型用于解答题 AI 批改

背景: 线上解答题批改的核心服务为一个基于 bert 的 QA 任务; 尝试基于 LLM 为代表的生成式人工智能代替以 bert 为代表的判别式人工智能

挑战: (a) 线上服务基于教育场景训练了自研 bert 预训练模型

(b) QA 任务数据集中的 startid 和 endid 为脚本自动生成, 标签质量较差

行动: 针对挑战 (a), 模型上尝试基于当前社区氛围较好、性能位于第一梯队的 Baichuan2-7B-Chat、Baichuan2-13B-Chat、Qwen2-7B 等主流 LLM 进行 SFT

针对挑战 (b), 采用 1-2round 混合训练, 避免使用质量较低的 startid 和 endid 标签

结果: 模型层面 LLM 相比 Bert 在 ExtractMatch 指标上提升 1.82 个点; 业务指标上, 解答题批改指标提升 4 个点

上海交通大学科研助理期间

2024.09 – 至今

项目名称: 科研助理

项目时间: 2024.09 – 至今

工作内容细分:

(1) 负责: 独立从 0 到 1 研发 mini 版 OCR 引擎

背景: 与卢长胜博士合作, 研发极简版 OCR 引擎用于工业场景下智能应用

任务: 引擎拆解为有向目标检测、文字方向判断和文字识别 3 个子模块, 需研发对应训练库、推理库和集成库

行动: (a) 训练库研发: 有向目标检测以 yolov11 为训练框架; 文字方向判断为自研分类训练框架; 文字识别以 mmocr-crnn 为训练框架

(b) 推理和集成库研发: 从 0 到 1 以 c++ 语言实现检测、分类和 sequence2sequence 推理库以及集成库

(c) 开发流程建立: 利用 C++ 包管理工具 CONAN 实现跨代码仓库之间的高效开发和版本维护

结果：该引擎复用性较强，已经在两个实际工业场景落地，项目介绍见 [johnson-magic/Production](#)

(2) 负责：point-supervised 有向目标检测方向研究

背景：对于目标密集场景，point-supervised 有向目标检测因其较低的数据标注成本具有较高的研究价值

任务：由杨学教授指导，与另外三名低年级研究生负责该课题的研究

结果：(a) 在六个主流的 benchmark DOTA-v1.0/DOTA-v1.5/DOTA-v2.0/DIOR/STAR/RSAR 上，达到 SOTA

(b) 相关成果整理为 [Point2RBox-v3](#) ICLR2026 已接收，共同第一作者（排名第一）

技能

学习能力：持续 follow 论文，并与公司业务、产品结合

编程语言/深度学习框架：Python/c/c++/pytorch

算法领域：通用目标检测、文字检测、文本识别、公式识别、生成式模型

工具类：git、docker、conan、CMake