

钟晓键

广东 | 130-4202-5045 | xiaojian.zhong@outlook.com

[xiaojianzhong.github.io](https://github.com/xiaojianzhong) | [LeetCode](https://leetcode.com/xiaojianzhong) | [GitHub](https://github.com/xiaojianzhong) | [知乎](https://www.zhihu.com/people/xiaojianzhong)

求职意向：机器学习平台研发 / C++ 后端开发



教育经历

西安电子科技大学 211 人工智能 智能感知与图像理解重点实验室 硕士（保研） 2021/09 ~ 2024/06

- GPA：学院前 10%
- 奖项：优秀推免生奖学金、一等奖学金、优秀研究生
- 课程：机器学习理论与实践 (98)、人工智能创新实验 (93)、图像工程 (96)、计算机视觉 (91)

暨南大学 211 网络工程 信息科学技术学院 本科 2017/09 ~ 2021/06

- GPA：院系前 5% (3.88)
- 英语：CET-6 (545)
- 职位：学习委员
- 奖项：一等奖学金、金龙鱼创新人才奖学金、5A 优秀学子创业之星、优秀毕业论文、优秀学生、优秀毕业生
- 课程：C++ 程序设计 (96)、操作系统原理 (97)、计算机网络 (98)、数据库系统原理 (95)、Unix / Linux (93)、机器学习与神经网络 (100)、数据挖掘 (95)、数字图像处理 (92)、软件测试与质量保证 (100)

实习经历

蚂蚁 支付宝技术部 算法工程师 2023/06 ~ 2023/09

实习工作：支付宝 AI 计算平台的服务建设 & 业务支撑 & 技术优化 C++ Nginx Redis gRPC PyTorch Stable Diffusion

- 开发 C++ 高性能计算服务，实现高效模型推理，支撑「收款码语音输入」「AI 照相馆」「AIGC 宠物皮肤」三业务上线
- 协定算法侧与业务侧接口，实践全链路标准研发流程：需求分析 - 排期 - 开发 - 自测 - CR - 联调 - 提测 - 灰度 - 全量
- 基于 Redis + Lua 实现分布式限流，针对计算密集型服务，实现集群级、业务粒度的流量控制，单机限流作为降级策略
- 使用 gRPC + ProtoBuf 重构现有 AIGC 服务，打通 Nginx 与 PyTorch 跨语言进程间通信，统一 PyTorch 模型交付形式
- 探索 Airflow 下的任务编排，通过 TaskFlow API 序列化 DAG，串联底层计算服务，向上游业务提供 pipeline 自定义能力

字节跳动 Data 架构 系统架构研发工程师 2020/09 ~ 2020/11

实习工作：抖音 & TikTok 推荐系统的离线数据流建设 & 资源占用优化 Python Redis 推荐系统

- 理解抖音短视频排序策略（预测模型、价值模型）、推荐架构（召回、粗排、精排）、内容数据流（正排、倒排）
- 基于 Pyflame + tracemalloc 定位 TikTok Python 正排更新服务的内存占用问题，本地 LRU cache 大 key 分拆至 Redis

腾讯 PCG 运营开发工程师 2020/06 ~ 2020/09

实习工作：腾讯微视内容特征中台的协议移植 & 服务重构 & 性能优化 Golang HTTP tRPC MySQL 设计模式

- 针对 HTTP 协议的「视频版本查询」服务不符合标准化规范，使用 tRPC-Go 实现协议移植，有效改善上游调用体验
- 引入 XORM 代替 Golang 原生 database/sql 标准库，实现 MySQL 表语义化，构建简易 kv cache 降低数据库查询频率
- 遵循司内 Golang 代码规范，以 MVC 为主线重构「视频文件下载」服务，单测覆盖率达 96%+，实习生 CR 评级达 A
- 优化下载逻辑，跳过不必要的文件落盘，协程池 + 任务队列实现异步审计，线上 QPS 提高 62%，下载耗时降低 200ms

顺丰物流 GetU 集运 系统运维 2019/06 ~ 2019/12

实习工作：Linux 后端服务器的资源负载监控 & 异常告警配置 & 定时任务管理

进兴科技 货多多集运 全栈开发 2018/06 ~ 2019/06

实习工作：技术合伙人，基于 Node.js + MongoDB 从零开始搭建 C 端中港物流集运平台，累计支撑 10w+ 香港用户

专业能力

- 熟悉 C++ 与 Python 语言，包括 STL、OOP、内存管理、多线程编程、C++11 标准，了解 CMake、Bazel 等构建工具
- 数据结构与算法基础扎实，LeetCode Knight 全站前 3%，能够手写字典树、并查集、树状数组、线段树等高级数据结构
- 熟悉操作系统基本知识，掌握 Linux 常用命令；熟悉 TCP/IP 网络协议栈、HTTP、RPC，了解 Linux 进程间通信编程
- 熟悉 SQL 语法，了解数据库范式，熟悉 MySQL 索引、事务、锁、存储引擎等概念，熟悉 Redis 常用数据类型
- 熟悉 Docker 常用命令，了解 Kubernetes 中镜像、容器、Pod、Service、服务发现、负载均衡等概念
- 了解分布式系统中的 CAP、BASE 理论，2PC、3PC 分布式事务，Raft、Paxos 分布式一致性协议、一致性哈希算法
- 熟悉 PyTorch 深度学习框架，了解 DP / DDP 分布式训练原理，了解混合精度训练、梯度累积等 GPU 显存优化策略
- 了解线性回归、逻辑回归等机器学习算法，熟悉图像分类、语义分割等视觉任务以及 ResNet、DeepLab 等神经网络
- 了解轻量级模型、模型剪枝、模型量化、知识蒸馏等常用的压缩加速策略，部署过 ONNX、TensorRT 等模型推理框架
- 掌握 Git、Vim、VSCode、Markdown 等软件开发工具链，了解常用设计模式，熟悉工业界研发流程与规范

项目经历

- 基于 PyTorch 开发遥感语义分割框架 [rsi-semantic-segmentation](#)，支持 FCN、DeepLab、U-Net、PSPNet 等分割模型
- 鉴于常规模型在推理时不满足实时性，选用 Fast-SCNN、BiSeNet、LEDNet 等轻量级分割模型，高效提取病虫害区域
- 使用 trtexec、torch2trt 转换 TensorRT 模型格式，重写「自适应 2D 池化」算子，实现 C++ 下的 PyTorch 模型推理
- 尝试「后训练量化」「L1 剪枝」等压缩加速策略，部署至边缘设备 Nvidia Xavier NX，实现 1080p 下 15fps 弱实时推理
- 在大疆 Linux C++ SDK 上进行二次开发，将无人机视频流与模型推理结果回传至地面端，实现用户侧秒级可视化

基于边缘端 AI 推理的海战实时制导模型PyTorchONNX分布式训练西安市兵工研究所2022/09 ~ 2023/04

- 鉴于视频切帧后训练集规模庞大，使用 PyTorch DP / DDP 分布式训练工具，实现两台双卡 3090ti 服务器上的单机多卡 / 多机多卡数据并行，缩短单次训练时间至两天内，显著提升模型训练速度与 GPU 利用率
- 针对航拍图像分辨率过大，采取混合精度训练、原地算子、梯度累积等 GPU 显存优化策略，有效扩增单机 batch size
- 观察到坦克与车辆类间差距过小，采取 CutMix 离线数据增强策略加入大量训练负样本，目标误检率降低 30%+
- 顺应国产化要求，使用 onnx、onnxruntime、onnx-simplifier、netron 实现 PyTorch、ONNX、OM 间的模型格式转换，适配国产昇腾 310 芯片，部署至弹载华为嵌入式设备 Altas 200 DK

手机端摄像 emoji 识别游戏Emoji HunterTensorFlow.js图像分类课程大作业（组长）2020/03 ~ 2020/06

- 实验量化对比 VGG、ResNet、DenseNet、Xception 以及 SqueezeNet、MobileNet 等（轻量级）图像分类模型
- 使用网格搜索策略调优学习率、宽度乘子等超参数，采取 CAM、Grad-CAM 可视化 CNN 特征图，以此验证分类能力
- 基于 TensorFlow.js 将分类模型部署至手机端浏览器，通过 webcam API 采集后置摄像头画面，模型实时推理识别 emoji
- 针对嘈杂背景对分类精度的影响，调用百度图像主体识别 API，裁剪主体区域以忽略背景，一定程度上降低误识别率
- 撰写 emoji 数据集构建指南，搭建图片收集平台，引导 4 名组员在 7 天内高效收集了 60+ 个 emoji 的 60k+ 张网络图片

竞赛经历

顺丰 SF-X 智慧供应链高校菁英编程赛

全国前 50 名

2023/06

LCCUP 力扣杯春季编程大赛

全国前 5%

2023/04

华为全球校园 AI 算法精英大赛 - [车道渲染数据智能质检](#)

2022/06

- 使用 ResNet、EfficientNet、Inception 等图像分类模型识别渲染图片缺陷，针对性设计水平翻转、随机旋转等数据增强
- 观察到缺陷类别分布不一致，在训练阶段采取 Focal Loss、类别平衡采样、困难样本挖掘 OHEM 等策略
- 在推理阶段使用投票融合、均值融合等模型融合方法复用高分模型，采取测试时增强 TTA 有效优化输出分割图质量
- 鉴于有监督训练中未利用 130k+ 张无标签图像，设计基于自训练的半监督算法，训练时为无标签样本动态生成伪标签

中科星图杯 - 海洋一号可见光图像中海冰语义分割

2021/11

- 针对大分辨率采取有重叠滑窗裁切方法制作训练集，使用 DeepLab v3+、PointRend 等语义分割模型划定海冰区域
- 观察到正负样本像素数量比例约为 1:4，尝试过采样正样本图像重置训练集，并使用 IoU 友好的 Dice Loss
- 观察到推理结果中海冰边缘表现较差，加入基于二分类交叉熵的边缘检测分支作为辅助任务，有效提高海冰边缘精度

商汤 OpenMMLab 算法生态挑战赛

季军

2021/09

- 为 MMSegmentation 语义分割框架加入 ECCV 2020 论文 [SFNet](#) 复现支持，各 backbone 下精度均达到原始论文水平

蓝桥杯全国软件和信息技术专业人才大赛 - Java 个人赛

广东省二等奖

2020/10

中国高校计算机大赛 - 网络技术挑战赛

华南赛区一等奖

2020/09

腾讯杯全国大学生人机 debug 测试大赛

全国第一名

2019/12

全国大学生软件测试大赛

全国第十二名

2019/12

「赢在创新」暨南大学创新大赛

校一等奖 + 最佳项目奖

2018/06

「挑战杯·创青春」中国大学生创业计划竞赛

广东省铜奖

2018/05

暨南大学 ACM 程序设计大赛

校第五名

2018/04

科研经历

M³DGCN: Dynamic Graph Convolutional Network with Multi-Teacher Multi-Objective Meta-Learning for Zero-Shot Hyperspectral Band Selection

论文

2023/06

- 将通道及其关系建模为图的节点和边，构建动态图卷积网络挖掘通道间共用特征，将网络结构与通道数量解耦
- 设计多教师知识蒸馏结构为特征选择网络提供蒸馏损失，引入优化多样性，提高跨数据集知识泛化能力
- 定义基于不确定性的多任务学习损失，联合优化分类与蒸馏任务，为两项损失自动、动态、并行地分配不同权重

基于上下文信息聚合和辅助学习的轻量级语义分割方法及系统

专利

2021/06

Awesome Weakly Supervised Semantic Segmentation

个人项目

2023/04

- 调研并汇总近 10 年弱监督语义分割领域 105 篇顶会预刊论文，[GitHub 仓库](#)获 100+ stars，[知乎专栏](#)获 700+ 收藏