司玉栋

\(\) 19121726080

https://github.com/ydongs

○ 教育经历

同济大学 - 硕士 - 集成电路工程专业

南通大学 - 本科 - 电子信息工程专业

2019.09 - 2022.03

2014.09 - 2018.06

卧工作经验

AI框架工程师(Triton编译器)@英特尔亚太研发有限公司

2022.07 - 2025.07

Triton 编译器 XPU 后端开发

- 功能实现
 - ・ 增加GLM支持、AOT编译、TF32的OpenCL扩展等开发
 - ・ 构建编译器后处理优化管道 (Postprocess Pass)
- 性能调优
 - 建立首套Triton-XPU性能Benchmark体系(覆盖Softmax, GEMM, FA),支持灵活拓展手写库(如XeTLA, CUTLASS, oneDNN)用于性能对比
 - · 性能优化与提升,关键内核性能达到Intel性能库XeTLA的90%+
- BUG修复
 - · 累计解决编译器后端 35+ High issue

PyTorch 生态优化

- ・ CI/CD 创新
 - 针对 PyTorch Inductor CPU Performance 基于 Jenkins 设计 AWS Xeon 实例自动化流水线,实现:
 - 自动收集200+模型的性能指标
 - 自动化报告生成与发布

CUDA测开实习生@英伟达半导体科技(上海)有限公司

2021.08 - 2021.11

CUDA Orin Simulator实践

- ・ 参与NVIDIA Orin t23x SOC的CUDA安全与代码覆盖测试体系构建
 - · 基于VDK虚拟测试套件搭建自动化测试平台
 - · 定位failed/timeout用例
 - 实现测试镜像版本自动化追踪

≫ 项目经历

驾驶员疲劳检测系统 | 计算机视觉 | 2021.07 | 校园项目

提出眼口姿态+头部欧拉角的多特征融合检测模型

- 设计眨眼计数(EAR<0.25)与哈欠检测(MAR>0.8)联合判据
- 在开发板部署轻量化推理管道
- · 开发QT监控界面,支持疲劳状态可视化告警

汽车氛围灯触控系统 | 嵌入式开发 | 2020.02 | 校园项目

开发车规级可编程氛围灯控制系统,实现触摸屏到LIN总线的全链路控制

- · 设计组态屏UART指令解析引擎,支持自定义控制指令
- · 实现LIN总线无条件帧传输协议
- · 开发电容触摸HMI界面

♡ 专业技能

- ・ 语言: C/C++, Python, Bash
- 基础:数据结构与算法,操作系统,计算机体系结构
- AI基础设施: PyTorch,LLVM,MLIR,oneAPI,OpenCL,CUDA,SYCL,Jenkins,GitHub Actions,Docker

