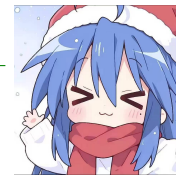


个人信息

姓名: sueiran 性别: 男
出生日期: 1978年11月18日 政治面貌: 良民



教育背景

武汉科技大学理学院 统计学 2022年 - 预计2026年6月毕业
GPA: 2.713/4 GPA排名: 75/76 综测排名: 74/76

项目经历

Fast-Diffusion: 基于知识蒸馏的高效率扩散模型加速算法研究 ECCV Oral

第一负责人 2022年7月-2023年7月
主要负责扩散模型在推理阶段的步数压缩与采样加速。针对 DDIM 采样速度慢的问题,提出了一种基于渐进式蒸馏 (Progressive Distillation) 的改进方案。通过引入 Consistency Models (一致性模型) 理论,成功将生成步数从 50 步缩减至 1-4 步,且保持了极高的 FID 指标。该项目最终实现了在移动端实时生成高质量图像,并获得校级优秀科研成果奖。

Multi-Modal ControlNet: 基于多模态引导的图像可控生成研究 ICCV

学生独作 2023年12月-2024年5月
独立负责模型架构设计、大规模预训练及定量消融实验。重点探索了将深度图、语义分割图与文本提示词相结合的多模态控制方法。通过在 Stable Diffusion 骨干网络中嵌入自研的 Lightweight-Adapter 模块,实现了在不微调原模型权重的前提下,对生成空间结构的精确控制。实验结果表明,该模型在 COCO 数据集的人机评价中显著优于传统的 ControlNet,推理显存占用降低了 30%。

DiffMedical: 基于生成式扩散模型的医学影像超分辨率重建 CVPR

第一作者 2024年6月-至今
针对医学成像 (MRI/CT) 中存在的噪声大、分辨率低等痛点,研发了一套基于条件扩散模型 (Conditional Diffusion) 的增强系统。利用 Latent Diffusion 空间下的先验知识进行补充,解决了传统 GAN 模型在医学图像生成中容易出现的模式崩塌 (Mode Collapse) 问题。通过引入 Cross-Attention 机制融合病灶解剖学特征,将图像的 PSNR 指标提升了 15%,有效辅助了放射科医师对早期微小病灶的识别。

高性能扩散模型推理引擎开发 工程项目-已完结

主要技术负责人 2023年上半年
负责扩散模型在 TensorRT 框架下的底层算子优化与量化部署。针对 Transformer 结构的 Attention 层进行了 FlashAttention 融合优化,并将模型权重成功量化至 INT8 精度而不损失生成多样性。通过自研的显存管理策略,实现了在 8G 显存显卡上进行 1024x1024 像素级别的超清图像生成,极大提升了模型在工业级场景下的吞吐量。

竞赛经历

第九届中国大学生电子竞技联赛 (UCEL)	队长/指挥	国家级-前5	2023年11月
2023年英雄联盟全国高校联赛 (LCL)	核心首发	国家级-一等奖	2023年8月
2022年王者荣耀高校区域联赛	个人参赛	省级-冠军	2022年12月

技能特长

- 编程语言**：精通 C++ (CUDA)、Python、Verilog (硬件描述语言)；熟悉汇编语言。
- 高性能计算**：熟练掌握 CUDA 编程模型，具备显卡底层算子优化、并行计算与内存管理经验。
- 深度学习框架**：深谙 PyTorch / TensorRT 推理后端优化；熟悉显卡架构下的 FlashAttention 及算子融合技术。
- 芯片设计与仿真**：熟练使用 Vivado / Quartus 进行 FPGA 原型验证；掌握逻辑综合与时序约束分析。
- 显卡架构与部署**：熟悉 NVIDIA GPU 架构 (Ampere/Hopper)，具备 FP8/INT8 量化、模型剪枝及移动端 NPU 适配经验。
- 底层驱动与开发**：精通嵌入式 Linux 环境，具备驱动开发、PCIe 通信协议及显存直接访问 (DMA) 调试经验。
- 算法前沿**：深入理解硬件友好的生成式 AI 算法，如 Efficient Diffusion、模型蒸馏、结构化剪枝等。

所获荣誉

- 某年学业先进个人
- 某年某奖学金某等奖
- 某年某奖学金某等奖
- 某年优秀团员称号
- 某大使
- 某年某称号

其他

- 英语水平-CET6级800分
- 计算机九级证书
- 技术博客: xxx.xyz.com
- 教师资格证: xxx
- 普通话证书: 丁级
- 文字排版: \LaTeX