

# 赵宇



138-6507-7874



1599601858@qq.com



shanghaitech

## 教育经历

上海科技大学&&中国科学院大学		2023.09 – 至今
信息与通信工程	硕士	
安徽工业大学		2019.09 – 2023.06
软件工程	本科	
• GPA: 3.5/4.0		

## 工作经历

算法工程师	上海张江数学研究院	2024.03 – 至今
• C++ 服务器开发		
为了增强项目的整体安全性，将原本使用的SSL加密模块更新为采用自主开发的TLSv1.3协议。在开发过程中，网络通信模块主要使用C++14标准库进行封装，主要以更新boost库模块。分别完成了Windows和Android平台的TLSv1.3加密模块，便于项目顺利部署。		
• 三维重建		
基于colmap,openmvs,openODM完成卫星地图和航拍地图三维重建算法的开发，解决了二维正射影像到三维tiles的映射，此外，设计并调整图像增强网络SpA-GAN,CT-GAN，训练模型调优，优化卫星成像。		
• 自动化标注软件构建		
基于DB-Net，自定义DeformConv算子，并注册到推理框架中，实现图像中的文本定位；基于Parseq-Net，替换backbone,设计特点场景的模型训练调优，提高文本识别率；将两个模型convert到onnx格式，使用onnxruntime推理部署集成到自动标注化软件中。		
• AI模型Add-Muti-Net编译部署		
完成模型的前端解析AMN-to-MLIR，模型的后端Pass优化,Convert和Lowering，实现了模型在CPU端的SLP矢量化，在GPU端，我们对模型进行多个阶段的下降，分别在High-AMN,Low-AMN,最后下降到GPU-Dialect,在这个IR中，我们考虑了BufferDeallocation,CopyElimination优化;并利用Pybind11封装成PythonAPI调用。部署阶段，我们利用Triton-server封装接口以及基于Backend开发适配我们的算法模块。		

## 街景定位

我们基于CLIP设计了特定的网络模型，实现了地理坐标与图片的对齐。针对Image-Pipeline我们采用ViT进行特征提取，针对Location-Pipeline我们使用GaussianEncoding编码向量，再通过MLP提取特征，经过测试在上海地区1km内准确率达到95.1%。之后我们利用xfeat匹配1km以内的全部图片，以达到精准定位，同时我们考虑匹配的性能，我们将模型部署到Triton-server中，利用多实例推理提高效率。

## 科研经历

BeHavior Prediction	CCF-B	2023.09 – 2024.01
---------------------	-------	-------------------

我们实现了一个pipeline对游戏行为进行预测。该项目旨在使用DETR模型进行目标检测识别人物位置和类别，之后送入到ResNet中进行行为分类，最后将得到一序列的数据送入到时序模型LSTM中进行预测。这包括大量的数据预处理任务，例如数据清理、特征编码和序列化。在模型选择方面，我评估了LSTM、Transformer模型，以及自定义模型。我使用收集的数据训练了模型，并使用交叉验证等技术对其进行评估。该项目产生的论文已被IUI接收（B类会议）。

Exploring Personality-Driven Human-AI Collaboration	CCF-A (在投)	2024.11 – 2025.01
---	------------	-------------------

本文论创新性地将Big-Five性格模型融入狼人杀游戏Agent的设计中，通过双路径架构实现了性格驱动的智能决策系统。我们采用RL-LLM微调方法，将历史对话和游戏状态作为输入，通过结构化输出实现理性决策；同时结合Prompt Engineering技术，通过动态调整提示词引导Agent展现不同性格特质的行为表现。这种融合方法既保证了游戏策略的合理性，又能生动体现不同性格特征（如外向性、神经质等）对玩家行为模式的影响，实现了技术性与拟真性的平衡。该方案为基于性格模型的游戏AI设计提供了新思路，可扩展应用于社交模拟、心理研

究辅助等多个领域。

### 卫星系统抗辐射加固

2023.11 – 2025.07

为了使Linux操作系统能够在卫星上稳定运行，我们提出了加固Linux内核的办法以及结合硬件ECC应对辐射引起的SEE效应。主要为内核中的关键数据结构保护，我们基于LLVM实现了TMR和CFCSS，分别对重要模块和函数进行控制流校验和冗余备份执行，有效提高了系统抗辐射能力。

### 模型抗辐射加固

2024.11 – 2025.07

为了使模型能更好地抵抗空间中重粒子辐射导致的比特翻转效应，我们从模型参数和模型编译运行角度进行了加固。参数方面，我们对模型进行4-bit量化并进行噪声注入，让模型降低对比特翻转的敏感程度，另外编译层，我们考虑了使用torch-mlir,将模型lowering到MLIR层，我们实现一个校验Pass，针对于Conv和Linear操作前后插入该Pass。通过BFA容错注入对比，我们针对于传统未加固模型将准确率下降10%的情况下需要的flip数量提高了7.2倍。

## 技能

- 具有工程开发经验，大部分工作均在工程落地使用
- 熟练掌握Linux环境下的Docker、C++、Python、Java，CD/CI工具使用
- 有一定的模型训练推理加速经验，了解模型训练和推理卡框架如：deepspeed,Llama-cpp,optimum
- 熟悉TVM,XLA,triton 深度学习编译框架，对arm,mips,cpu0等指令集有一定了解
- 掌握使用MLIR ,LLVM 对于 Dialect,Pass优化的开发

## 其它

- Github:<https://github.com/GIN-ZHAO>
- Language: - CET6( score 460)