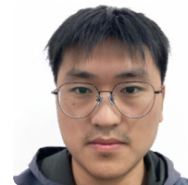


# 孙爱华

15396897213 | aeeeeeeep@proton.me

求职意向：算法工程师 | 期望城市：上海



## 个人优势

- 熟悉 CNN, RCNN, Fasttext, Transformer 等经典网络结构原理, 了解 Vit, CLIP, SwinT 等视觉Transformer网络
- 了解 GPU 并行架构, 在 CV 技术指南 公众号发表 CUDA 教程十余篇, 累积 4w 字
- 熟悉 Linux 环境下深度学习算法的开发, 并使用 Docker 部署项目
- 掌握 onnxruntime, TensorRT 系列部署框架的使用, 掌握 PyTorch 框架
- 长期编写技术博客, 全平台总访问量3.3w, 个人网站: <https://aejeeep.top>
- 乐于开源代码 Github 主页: <https://github.com/aejeeep>

## 教育经历

盐城师范学院 本科 数字媒体技术 2020-2024

GPA: 3.32/4

### 荣誉奖项

- 2022 微信大数据挑战赛 多模态短视频分类 三等奖 Rank 72  
(参赛代码: <https://github.com/aejeeep/2022WBDC-semi>)
- 2022 MathorCup高校数学建模挑战赛 全国二等奖
- 2021 APMCM亚太地区大学生数学建模竞赛 全国二等奖
- kaggle: RSNA 2022 Cervical Spine Fracture Detection TOP 24%
- kaggle: DFL - Bundesliga Data Shootout TOP 37%
- 2023 全球人工智能技术创新挑战赛 医学影像诊断报告生成 Rank 123  
(参赛代码: <https://github.com/aejeeep/2023GAIC>)

## 专业技能

Python, Shell, CUDA, OpenCV, C/C++, Docker

## 项目经历

基于多模态特征的大黄蜂入侵物种防治算法 算法开发 2023.03-2023.05

### 内容:

大黄蜂作为一种入侵性物种, 对生态系统和农业产生了严重的影响。为了有效防治大黄蜂的入侵, 本项目提出了一种基于多模态特征的防治算法。通过结合图像识别、用户文本描述和上报地点等多模态信息, 实现对大黄蜂的防治与繁殖预测, 为后续的防治工作提供可靠的依据。

### 业绩:

对于图像模态, 我们使用 Convnext 模型, 对收集到的大黄蜂图像进行了训练和分类。识别大黄蜂的存在与否。对于文本特征, 使用 Fasttext 对用户的文本描述进行情感分析和关键词提取, 再编码识别。最后使用熵权法研究不同特征对大黄蜂繁殖的重要程度, 进行加权判定, 在测试集上达到了 99% 的准确率。

难点: 样本分布极度不平衡, 通过重采样, 对抗训练, 提示词生成稀少类别文本样本缓解。采用 Bert 模型融合方案虽然可以更好融合特征, 但本项目文本过短, 有过拟合现象, 遂采用决策融合方案。

项目链接: <https://github.com/aejeeep/BumblebeeSightings>

电力生产环境员工安检系统 算法开发 2022.12-2023.03

### 内容:

基于现场监控摄像头视频进行智能分析, 作业前, 进行人员着装与绝缘防护判断, 对作业的配电柜与设备初始状态, 做时态信息和票面信息对比研判, 对错误的配电柜和错误初始状态及时告警。

### 业绩:

使用 Yolo 目标检测网络进行实时着装检测, 使用 OpenCV 传统视觉方法对设备状态进行检测。并对算法在 Nvidia 环境下的量化, k8s 集群上的部署, 实现 50 fps (96.8 acc) 的检测速度。设备检测算法实现 500 fps (99.5 acc) 的检测速度。

难点: Yolo 检测候选框重叠, 使用 NMS 合并。视频传输占用带宽过高, 且数据库压力过大, 使用跳帧检测, 并对检测结果编码压缩。

使用技术：CUDA, Tensorrt, nvidia-docker, OpenCV, Pytorch  
项目链接：<https://github.com/aeeeeep/ESDSalgo>

## 微型空气站浓度类监控动态校验系统      算法开发

2023.02-2023.02

### 内容:

对微型空气站检测数据浓度的数据偏移进行校正。

### 业绩:

使用机器学习 GBR 算法方法对微型空气站浓度类挖掘特征,对偏差进行校正。训练调参采用网格搜索,编写算法训练与预测以及模型管理的接口。

难点：噪声过多，使用小波变换和四分位滑动窗口减少噪声。随机采样预测精度低，使用时间序列采样。

使用技术：Sklearn, Pandas

项目链接：<https://github.com/aeeeeep/GBRalgo>

## 实习经历

### CV技术指南公众号

2022.09-至今

#### 内容:

在 CV 技术指南公众号负责编写 CUDA 专栏文章。发表 CUDA 教程十余篇,累积4w字。

#### 实习成果:

以发表的包含如下内容:

1. GPU 编程概述 和 CUDA 环境搭建
2. CUDA 模型概述
3. CUDA C 编程简介
4. PyCUDA 编程简介
5. CUDA 内存模型概述
6. CUDA 内存管理(一)
7. CUDA 内存管理(二)
8. CUDA 内存管理(三)
9. CUDA 流,事件与同步
10. CUDA 流,并发与上下文

后续会更新:

11. CUDA 代码的调试与性能分析
12. CUDA 性能优化
13. Python中 Scikit-CUDA 模块使用 CUDA库
14. CUDA 设备函数库
15. CUDA C 实现神经网络
16. CUDA Samples 代码阅读

## 资格证书

大学英语四级

## 社交主页

<https://github.com/aeeeeep>

<https://aeeeeep.top>

## 兴趣爱好

写技术博客，参加深度学习竞赛，阅读开源项目