

牛志康

(+86) 186-2861-8289 · zhikangniu@sjtu.edu.cn · GitHub @ZhikangNiu

教育背景

上海交通大学, 计算机科学与技术, 在读博士研究生 2024.9 - 2029.6
西安电子科技大学, 人工智能, 工学学士 2020.9 - 2024.6
GPA: 3.8/4.0, 排名 7/138(前 5%), 国家奖学金 (2021-2022 学年度, 排名 2/138), 校级一等奖学金 (2 次)

实习经历

微软亚洲研究院 | 自然语言计算组 (NLC Group), 算法实习生 2023.8-2024.8
在刘树杰博士和周龙博士的指导下, 探索适用于语音生成任务的语音离散化模型 (Audio Codec), 副语言可控的语音合成模型 (Text to Speech) 和多模态大模型 (Multi-Modality Large Language Model)
• 基于 fairseq 框架独立完成了语音离散化模型的训练、推理、评测代码。并基于该框架上探索鲁棒的语音离散化模型。产出一篇论文已投稿至 (SLT 2024) 2024 IEEE Spoken Language Technology Workshop。
• 调研现有的语音合成工作, 并构建了带有副语言标注的中文语音合成数据集, 基于 VALLE、MELLE 等 decoder-only 模型探索具有表现力和可控的语音合成模型。

项目经历

Thorough PyTorch (开源 PyTorch 教程) 负责人 2021.7 - 至今
• Thorough PyTorch 是一个开源的中文的 PyTorch 教程, 截至目前在 Github 获得 2300 余次 stars, 397 次 fork。获得人民邮电出版社计算机分部和上海浦视人工智能实验室 MMYOLO 开源算法库的支持。
• 贡献: 完成了教程内容整体的设计和大部分章节的撰写 (包括模型的构建、训练、测试、部署)
• 项目地址: <https://github.com/datawhalechina/thorough-pytorch>

High Fidelity Neural Audio Compression (论文复现) 负责人 2023.2 - 2023.9
• 在上海交通大学陈谐教授指导下, 在 LibriTTS 数据上复现了 EnCodec 的训练过程, 并在测试集上的测量的 PESQ、STOI、Mel Distance 和 STFT Distance 语音质量评价指标上优于开源的 EnCodec 模型, 目前代码和权重均已开源获得 126 次 stars, 并在社区的贡献下完成了双声道和 SIURM 训练。
• 贡献: 独立完成了 EnCodec 模型的训练代码和精度对齐 (包括单卡及分布式训练)
• 项目地址: <https://github.com/ZhikangNiu/encodec-pytorch>

荣耀-西电智能剪辑技术产品开发 (项目落地) 算法核心贡献者 2021.7 - 2022.10
• 在西安电子科技大学吴家骥教授指导下, 该应用对用户输入的视频进行自动剪辑, 实现智能挑选高光时刻、模板自动匹配、一键成片的功能。本项目已经集成至荣耀手机的 AI VLOG 视频大师应用, 首发于荣耀 Magic5, 现已全机型铺开。
• 贡献: 核心贡献者, 使用知识蒸馏的方法将 CLIP 模型中的视觉编码模型迁移至轻量化网络。基于 Place365 数据集训练场景分类模型, 并导出 ONNX 格式文件在手机端进行测试。
• 项目产出: 一篇国家发明专利在申 (第一作者) 和一项国家级大学生创新创业训练计划 (负责人), 中国国家发明专利已将技术交底书交付给荣耀公司。

科研论文

VALL-T: Decoder-Only Generative Transducer for Robust and Decoding-Controllable Text-to-Speech.
• Chenpeng Du, Yiwei Guo, Hankun Wang, Yifan Yang, **Zhikang Niu**, Shuai Wang, Hui Zhang, Xie Chen.
• Submitted **Conference on Neural Information Processing Systems (Neurips 2024)**
Fast-HuBERT: An Efficient Training Framework for Self-Supervised Speech Representation Learning.
• Guanrou Yang, Ziyang Ma, Zhisheng Zheng, Yakun Song, **Zhikang Niu**, Xie Chen*
• **IEEE Automatic Speech Recognition and Understanding Workshop (ASRU Workshop)**, 2023.
NDVQ: Robust Neural Audio Codec with Normal Distribution-Based Vector Quantization.
• **Zhikang Niu**, Sanyuan Chen, Long Zhou, Ziyang Ma, Xie Chen, Shujie Liu*
• **IEEE Spoken Language Technology Workshop (SLT)**, 2024.

竞赛获奖

• 第九届中国大学生“互联网+”创新创业竞赛, 全国铜奖	2023.9
• 美国大学生数学建模比赛, 国际一等奖 Meritorious Winner	2022.5
• 中国大学生计算机设计大赛, 西北地区一等奖	2023.5
• 国家级全国大学生创新创业训练计划 (视频帧质量分析与精彩集锦浓缩, 负责人)	2023.3
• 国家级全国大学生创新创业训练计划 (视频流智能封面推荐系统, 负责人)	2021.3
• 第七届全国大学生互联网加创新创业训练大赛, 陕西省金奖	2021.7
• 第八届全国大学生互联网加创新创业训练大赛, 陕西省铜奖	2022.7
• 全国大学生创青春挑战杯, 陕西省铜奖	2022.4

社区参与/实践其他

• 国内 AI 开源组织 Datawhale 成员	
• Datawhale 开源教程审核委员会成员	
• 西安电子科技大学-百度飞桨领航团团长	
• 西安电子科技大学-百度菁英班班长	