

夏一博 硕士二年级

出生日期: 2000.11.12

电话: (+86) 136 2471 1873

微信: ggsds6898

Email: yiboxia@buaa.edu.cn



教育经历

2023.1-至今 清华大学, 自动化系, 脑与科学认知研究院, 访问研究生, 导师: [刘烨斌教授](#)

2022.9-至今 北京航空航天大学, 宇航学院, 模式识别与智能系统, 硕士在读, GPA: 3.82/4.0
(以专业第一的成绩考入, 初试成绩392分)

2018.9-2022.6 北京航空航天大学, 宇航学院, 飞行器控制与信息工程, 工学学士, GPA: 3.53/4.0

研究方向

简述: 我硕士期间的主要研究方向是计算机视觉和图形学, 面向 2D/3D 数字人驱动和生成。具体来讲, 我聚焦于语音驱动的情感数字人脸生成, 同时对语音/文本驱动的手势、手语生成进行研究; 此外, 我还关注于高光谱图像的分类和生成。

2023.1-2024.5 脑与科学认知实验室, 自动化系, 清华大学, 由刘烨斌教授指导

- **语音驱动的情感数字人脸生成:** 提出了基于高斯混合模型的语音驱动情感视频肖像生成框架, 可根据语音生成指定情绪的高真实感说话人像视频, 并且可以在已知基本情绪状态之间实现自由、平滑的情绪过渡。相关工作拟投稿到 IEEE T-PAMI 期刊 (SCI, Q1, IF=23.6)

2022.1-2022.12 航天信息工程系, 宇航学院, 北京航空航天大学, 由罗晓燕副教授指导

- **少样本下高光谱图像的细粒度分类:** 开发了半监督的光谱-空间特征学习框架, 设计了用于多尺度的光谱-空间特征学习的粒度感知的折叠光谱的空间超分模块, 并利用无标注样本提升少样本情况下模型的鲁棒性。该工作已投稿到 Knowledge-Based Systems (SCI, Q1, IF=8.8)
- **基于 Diffusion 的高光谱图像生成:** 构建了基于 Diffusion 模型的 RGB 到高光谱图像的生成方案, 并成功在高光谱数据集上表现 SOTA。相关工作正在完善和整理, 拟投稿 SCI 期刊

学术论文

1. **Xia, Y.**, Wang, L., Deng, X., Luo, X., & Liu, Y. (2024). GMTalker: Gaussian Mixture based Audio-Driven Emotional talking video Portraits. IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI), 2024. (在投)
2. **Xia, Y.***, Zhang, L*., Luo, X., Li, S., Hu, Y., Fu, Y. (2024), S2GS2: Semi-Supervised Granularity-Aware Spectral-Spatial Feature Learning Framework for Fine-Grained Hyperspectral Image Classification. Knowledge-Based Systems. (Under Review)
3. Wang, Y., Liu, Z., **Xia, Y.**, Zhu, C., & Zhao, D. (2021). Spatiotemporal module for video saliency prediction based on self-attention. Image and Vision Computing, 112, 104216. <https://doi.org/10.1016/j.imavis.2021.104216>

项目经历

2023.8-2023.10 语音驱动的高真实感数字人脸生成和对话

主要负责人

(在清华大学自动化系刘烨斌教授和王立桢博士后指导下完成)

- 构建了高斯混合分布的连续、解耦的情绪隐空间, 并结合语音预测 3DMM 表情参数序列; 提出基于正则化流为先验 VAE, 根据音频自回归的预测头部姿态, 眨眼和眼球参数; 这些参数用于 NeRF/StyleGAN 的渲染, 以生成情绪可控、姿态可控的高真实感数字人视频
- 结合 ChatGPT, ASR, TTS, 声音克隆等开源技术, 可以用于虚拟人对话生成

2022.8-2023.4	一种基于视觉和船状态结合的无人机着舰状态监视方法	第一学生负责人
<ul style="list-style-type: none">➤ 采用目标跟踪和传统测距结合的方法估计无人机相对位置，采用 6DoF 姿态估计算法估计相对姿态，并考虑船的运动状态动态规划降落轨迹曲线，实现无人机的状态监视和控制➤ 已授权国家发明专利（专利号：ZL 2023 1 0113156.X）		
2024.1-2024.5	高辐照条件下机场大范围视觉增强技术，国家重点研发计划项目	主要负责人
<ul style="list-style-type: none">➤ 与成飞 611 所，聚集“综合导航多源信息设计验证环境技术开发”子课题➤ 具体负责大差异数据条件下的海上目标检测，小样本条件下的海上目标查印证		
2023.9	第一届智慧识别算法挑战赛——目标识别赛道	主要负责人
<ul style="list-style-type: none">➤ 基于改进的 YOLO-v8 网络，对虚拟 RGB 和红外图像进行小样本舰船检测		第 4 名
2023.7-2023.8	空军算法挑战赛	主要负责人
<ul style="list-style-type: none">➤ 基于传统机器学习算法和神经网络解决无人机蓝牙信号的分选和分类问题		第 20 名
2023.9-2024.5	iVoice+: 关爱残障人士的唇、手语一体化数字伴侣	直接+指导参与
<ul style="list-style-type: none">➤ 基于上述语音驱动数字人脸项目，指导本科生实现文本驱动的手语生成，构建手部精细化的半身 3D 数据集，实现聋哑人和正常人的无障碍沟通，还可用于聋哑人的教育、导游、医疗等。该项目已参加北航第 34 届冯如杯科技竞赛，相关论文拟投稿 SCI 期刊		
2024.1-2024.3	2 项国家自然科学基金的立项申请	参与
<i>(在清华大学刘烨斌教授和北京师范大学张鸿文副教授的指导下完成)</i>		
<ul style="list-style-type: none">➤ 主要负责“细腻表情和全身运动感知和捕捉”和“物理合规的单目全身精细运动捕捉”两章节的申请材料撰写，受到上述情感可控数字人工作的启发，用情感促进精细化的动作捕捉		

实习经历

2023.1-2024.5	脑与科学认知实验室，自动化系，清华大学，由刘烨斌教授指导
<ul style="list-style-type: none">➤ 完成语音驱动的情感数字人脸生成项目，并完成 T-PAMI 文章的撰写➤ 探索具身智能相关的研究；复现开源的 Mobile-Aloha 项目，目前仍在探索阶段	
2021.7-2021.8	浙江未来技术研究院 (嘉兴)，暑期实习
<ul style="list-style-type: none">➤ 完成“一种雾天低能见度标准供配电台区目标检测算法”项目，并获得优秀实习生称号	

荣誉奖项

2023	北航三好学生，北航研究生学业奖学金一等奖
2022	宇航学院优秀毕业论文，北航研究生学业奖学金一等奖
2021	北航冯如杯科技竞赛二等奖，美国大学生数学建模竞赛 M 奖
2020	北京市数学建模竞赛一等奖，北航学科竞赛奖学金一等奖
2019	北航优秀学生干部，北航社会工作奖学金一等奖
2017	全国中学生物理竞赛一等奖 (省级)

个人技能

<ul style="list-style-type: none">➤ 专业技能：Python, Pytorch, Matlab, C/C++.➤ 审稿人：2024 IEEE TVCG, 2023 CAAI➤ 兴趣爱好：篮球，排球，乒乓球。曾获得北航篮球联赛季军(3 次)，获得北航排球联赛亚军➤ 其他：获得北航数学二学位，英语六级	
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--