

# 左育莘

手机 : (+86) 13630049521 · 邮箱 : zuoyushen12@gmail.com · 主页: [个人主页](#)

## 教育背景

澳大利亚国立大学, 计算机学院, 博士生	2025.09 - 至今
• 研究方向: 图像 / 视频生成, 多模态大模型, AI 智能体	
清华大学, 自动化系, 硕士	2019.09 - 2022.06
• GPA: 3.58 / 4.0	
• 研究方向: 图像复原, 目标检测及分割	
西安电子科技大学, 电子信息工程, 本科	2015.08 - 2019.06
• GPA: 3.90 / 4.0 (Top 5%), 优秀毕业生	

## 研究经历

<b>4KAgent : 基于智能体 AI 的任意图像 4K 超分辨率系统 - NeurIPS 2025</b>	2025.01 - 至今
• 提出基于智能体 AI 的系统处理复杂图像复原任务, 并将图像提升至 4K 分辨率	
• 设计基于多专家混合 (Q-MoE) 策略的多智能体系统, 分析图像并提出执行计划, 提升复原质量	
• 构建专用人脸复原管道, 进一步增强图像中人脸细节	
• 开发配置模块, 支持用户定制系统以满足多样化复原需求	
• 在自然图像、AIGC 图像、遥感图像及生物医学图像的 4K 超分辨率处理中实现卓越性能	
• 合作者: 涂正中教授, 德州农工大学 TACO 团队。项目官网: <a href="https://4kagent.github.io">4kagent.github.io</a>	

<b>视觉语言模型对高斯噪声的鲁棒性研究及防御 - ICCV 2025</b>	2024.07 - 2025.03
• 首次系统性地进行分析, 揭示主流视觉语言模型缺乏对高斯噪声视觉扰动的鲁棒性	
• 提出 Robust-VLGuard 数据集对视觉语言模型进行微调, 在保持视觉语言模型通用性能的前提下提高模型对高斯噪声的鲁棒性	
• 将防御范围扩展至视觉对抗攻击, 提出结合扩散技术的防御框架 DiffPure-VLM: 将对抗噪声转化为类高斯噪声, 再利用通过 Robust-VLGuard 数据集微调后的视觉语言模型进行有效防御	

<b>基于一步扩散的多视角一致性图像风格迁移 - ECCVW 2024</b>	2024.06 - 2024.08
• 聚焦于三维场景中多视角图像的风格迁移, 提出基于一步扩散模型的图像风格迁移方法 OSDiffST	
• 利用 LoRA 技术实现预训练扩散模型在风格迁移任务中的快速适配, 提出视觉条件模块, 高效提取并注入目标风格图像的风格信息	
• 研究论文被 AI for Visual Arts Workshop and Challenges (AI4VA) 录用, 收录于 ECCV 2024	

<b>用于图像复原的 Visual Token Transformer</b>	2020.05 - 2021.06
• 基于 Visual Token, 设计 Transformer 模块来提取自然图像中的非局部自相似性以及跨尺度自相似性	
• 将 Transformer 计算复杂度从 $O(n^2)$ 降至 $O(n)$ , 同时保持相似图像复原性能	
• 方法收录在 NTIRE 2021 Challenge on Image Deblurring 文章中 ( <b>CVPRW 2021</b> )	

<b>低分辨率掌纹图像去噪 - Neurocomputing 2019</b>	2019.01 - 2019.06
• 设计基于生成对抗网络 (GAN) 的模型来解决掌纹图像中的多种类型的噪声, 并在训练中使用 Gabor 损失函数来保留更多的方向信息	
• 在测试数据集上, 所提方法在图像去噪质量和掌纹识别精度方面均优于现有的 state-of-the-art 方法	

## 实习经历

<b>微软亚洲研究院, 研究实习生</b>	2021.07 - 2022.07
• 旋转目标检测 (多方向表格检测) <ul style="list-style-type: none"><li>- 设计用于旋转目标检测的 anchor-free two-stage detector</li><li>- 设计用于旋转目标检测训练的序列不变损失 (sequence-invariance loss) 和相对偏移 (relative-offset)</li></ul>	

- 在产品数据集中，目标检测器在不同旋转角度下性能稳定 (F-score 波动  $\leq 0.02$ )
- 在产品数据集中达到 state-of-the-art 性能 (F-score) 并融合到 Azure OCR API 中
- 获得微软亚洲研究院实习生项目“明日之星”奖项

## 腾讯优图实验室，研究实习生

2020.10 - 2021.05

- **UniInst: Detection free and NMS free 实例分割 - CN114332457A [P]**

- Instance-aware One-to-one Assignment: 使用匈牙利匹配，根据分类分数和 Mask IOU 将最佳匹配特征点分配给目标作为正例点
- MaskIOU Branch: 在训练过程中，学习预测生成的 Mask 的 IOU，在推理过程中，将 Mask 的 IOU 预测与分类分数相乘作为最终的置信度
- 在 COCO test-dev 2017 和 OCHuman 数据集上取得 state-of-the-art mask AP

## 职业经历

### 香港理工大学（PolyU），Research Assistant

2024.04 - 2025.05

- **人工智能与信号处理实验室**

- 用于图像处理任务的扩散模型的加速（例如图像风格迁移任务 (AI4VA@ECCV2024)）
- 图像与视频超分辨率
  - \* 在 **CVPR 2025 NTIRE Challenge on Short-form UGC Image Super-Resolution** 获得第一名
  - \* 在 **ECCV 2024 AIM Challenge on Efficient Video Super-Resolution** 获得第二名

### 微软亚洲工程院，Applied Scientist in Bing

2022.08 - 2024.03

- **Bing News - 推荐系统**

- 动态配额分配
  - \* 训练分类器识别用户触发请求，并调整 Ranker 中各召回路径的配额
  - \* 产品性能：在保持用户体验表现的前提下降低约 20% 的计算资源消耗

- **Bing Whole Page - 大型语言模型应用**

- Bing 搜索中的 Answer triggering - 房地产相关
  - \* 利用 GPT-3.5 标注网页搜索结果中的高难度样本，获取 130 万个新训练样本
  - \* 基于通过大型语言模型标注增强后的新训练集，训练 Answer trigger 模型
  - \* 产品成效：Bing 房地产垂直领域的 DAU（每日活跃用户）增加了 4.1K

## 发表文章

- **Yushen Zuo, Qi Zheng, Mingyang Wu, Xinrui Jiang, Renjie Li, Jian Wang, Yide Zhang, Gengchen Mai, Lihong V. Wang, James Zou, Xiaoyu Wang, Ming-Hsuan Yang, Zhengzhong Tu.** “4KAgent: Agentic Any Image to 4K Super-Resolution”, **NeurIPS 2025**.
- Jiawei Wang\*, **Yushen Zuo\***, Yuanjun Chai, Zhendong Liu, Yicheng Fu, Yichun Feng, Kin-Man Lam. “Safe-guarding Vision-Language Models: Mitigating Vulnerabilities to Gaussian Noise in Perturbation-based Attacks”, **ICCV 2025**. (\*: Equal Contributions)
- **Yushen Zuo, Jun Xiao, Kin-Chung Chan, Rongkang Dong, Cuixin Yang, Zongqi He, Hao Xie, Kin-Man Lam.** “Towards Multi-View Consistent Style Transfer with One-Step Diffusion via Vision Conditioning”, **ECCV 2024 Workshop**.
- Zongqi He, Zhe Xiao, Kin-Chung Chan, **Yushen Zuo**, Jun Xiao, Kin-Man Lam. “See In Detail: Enhancing Sparse-view 3D Gaussian Splatting with Local Depth and Semantic Regularization”, **ICASSP 2025**.
- Xin Li, Kun Yuan, Bingchen Li, Radu Timofte, **Yushen Zuo** et al. “NTIRE 2025 Challenge on Short-form UGC Video Quality Assessment and Enhancement: Methods and Results”, **CVPR 2025 Workshop** (1st place).
- Marcos V. Conde, Zhijun Lei, Wen Li, Christos Bampis, **Yushen Zuo** et al. “AIM 2024 Challenge on Efficient Video Super-Resolution for AV1 Compressed Content”, **ECCV 2024 Workshop**.
- Seungjun Nah, Sanghyun Son, Suyoung Lee, Radu Timofte, Kyoung Mu Lee, **Yushen Zuo** et al. “NTIRE 2021 Challenge on Image Deblurring”, **CVPR 2021 Workshop**.
- Shengjie Chen, Shuo Chen, Zhenhua Guo, **Yushen Zuo**. “Low-resolution palmprint image denoising by generative adversarial networks”, **Neurocomputing 2019**.

## 荣誉奖项

---

CVPR 2025 NTIRE Challenge on Short-form UGC Image Super-Resolution - 第一名	2025.03
AIM 2024 Challenge on Efficient Video Super-Resolution - 第二名	2024.08
微软亚洲研究院实习项目“明日之星”奖	2022.05
CVPR 2021 NTIRE Image Deblurring Challenge - Track1. Low Resolution (10 / 60)	2021.03
Kaggle NFL 1st and Future - Impact Detection: 银牌 (23 / 459)	2021.01
第一届海洋目标检测国际挑战赛冠军 (1 / 151)	2020.12
校级优秀毕业生	2019.06
美国大学生数学建模竞赛 (MCM/ICM) M 奖 (Meritorious Winner)	2018.05
第八届全国大学生数学竞赛 (非数学类) 陕西赛区一等奖	2017.08
2016, 2017, 2018 学年校一等奖学金, 校优秀学生	

## 技能

---

编程语言 Python (PyTorch, NumPy, Scikit learn. etc.), C/C++, HTML/CSS, SQL.  
其他 Linux, Shell (Bash/Zsh), LATEX(Overleaf/Markdown), Microsoft Office, Git.