

王帅

南京大学计算机科学与技术系 直博生

wangshuai.mcg@smail.nju.edu.cn · GitHub: WANGSSSSSSS

个人总结

本人乐观向上，工作负责，自我驱动力强，热爱尝试新事物，同时对自己有清醒认识。研究兴趣包括生成模型的可扩展训练与快速采样，尤其关注扩散模型、视觉生成、视觉感知与表征学习。

教育背景

南京大学，计算机科学与技术系	2022.09 – 至今
计算机科学与技术，直博生在读	
西北工业大学，计算机学院	2018.09 – 2022.06
计算机科学与技术，本科	

竞赛获奖与荣誉

- 2018–2019 年度 国家奖学金
- 2018–2019 年度 校级优秀学生
- 2019 FIRA RoboWorld Cup RoboSot Race (Development Group) 亚军
- 2019–2020 年度 吴亚军奖学金
- 2019–2020 年度 校级优秀学生
- 2019–2020 年度 全国大学生数学竞赛（非数学类，陕西赛区）一等奖
- 2020 ICRA RoboMaster AI Challenge 国际三等奖
- 2021 火箭军“智箭-火眼”人工智能挑战赛“珠联璧合”科目 冠军
- 2022–2023 年度 一等学业奖学金
- 2022–2023 年度 腾讯奖学金（硕士）
- 2023–2024 年度 一等学业奖学金
- 2024–2025 年度 国家奖学金（博士）

发表论文（一作）

Deep Equilibrium Object Detection Shuai Wang, Yao Teng, Limin Wang	ICCV 2023
• 针对当前 query-based 目标检测器参数冗余、性能未饱和的问题展开研究。 • 将 FFN 检测器 decoder 转化为 RNN 形式以减少参数冗余，并通过 DEQ 模型进一步探索性能饱和上限。	
Deep Equilibrium Object Detection and Segmentation Shuai Wang, Yao Teng, Limin Wang	TPAMI
• DEQDet 的期刊扩展版本，增加实例分割应用，并补充更多检测与分割实验。	
FlowDCN: Exploring DCN-like Architectures for Fast Image Generation with Arbitrary Resolution Shuai Wang, Zexian Li, Tianhui Song, Xubin Li, Tiezheng Ge, Bo Zheng, Limin Wang	NeurIPS 2024
• 提出 Groupwise-MultiScale-DCN 作为基础构建块，探索 DCN 结构在图像生成中的应用。 • 相比 DiT/SiT，参数量减少 8%，FLOPs 减少 20%，训练仅需五分之一总图片数。 • 在 ImageNet 256 x 256 上取得 FID 2.00 和 SoTA sFID 4.35。 • 在 ImageNet 512 x 512 上取得 FID 2.44 和 SoTA sFID 4.53。	
Differentiable Solver Search for Fast Diffusion Sampling Shuai Wang, Zexian Li, Tianhui Song, Xubin Li, Tiezheng Ge, Bo Zheng, Limin Wang	ICML 2025
• 使用极少参数量搜索具有泛化性的 solver 参数。 • 对于 FM 模型，SiT 的 10 步 FID 提升到 2.40，FlowDCN 提升到 2.35，显著超越现有 solver。 • 对于 DDPM 模型，DiT 的 10 步 FID 提升到 2.33，显著超越 DPM-Solver++/UniPC。	
DDT: Decoupled Diffusion Transformer Shuai Wang, Zhi Tian, Weilin Huang, Limin Wang	ICLR 2026
• 解耦生成模型预训练中的高频解码与低频编码。 • 在 ImageNet 256 x 256 上取得目前最强 FID 1.26，公平对比下领先其他算法。 • 在 ImageNet 512 x 512 上取得目前最强 FID 1.28，大幅领先其他算法。	
PixNerd: Pixel Neural Field Diffusion Shuai Wang, Ziteng Gao, Chenhui Zhu, Weilin Huang, Limin Wang	CVPR 2026
• 利用神经场辅助大像素 patch 解码。 • C2I 模型在 ImageNet 256 x 256 上取得 FID 1.78；相同计算量下接近 latent 模型性能，并大幅领先其他 pixel 生成模型。 • T2I 模型取得 GenEval 0.73 和 DPG 80.9，接近部分先进文生图模型，并大幅领先像素生成模型。	
UniDDT: Unifying Multimodal Understanding and Generation with Decoupled Diffusion Transformer Shuai Wang, et al.	ECCV 2026
• 基于 Decoupled Diffusion Transformer 统一多模态理解与生成。	

- 验证 DDT 类结构在不同视觉表征空间中的可扩展性。

项目经历

西北工业大学足球机器人基地 – FireFly	2019.10 – 2021.01
• 负责 RoboMaster 人工智能挑战赛中的实时目标检测模块。 • 负责 RoboMaster 人工智能挑战赛中基于哨岗相机的全场定位与目标检测模块。	
西北工业大学三维视觉实验室	2020.09 – 2021.01
• 参加火箭军人工智能挑战赛，获得精度第一，速度领先第二名一倍。 • 参与基于光流方法 RAFT 的双目视差估计科研项目。	
商汤科技	2021.12 – 2022.05
• 参与商汤绝影自动驾驶感知任务。	
阿里妈妈	2024.04 – 2024.12
• 参与智能创作中的模型加速，与正式员工共同研发 Flux-Turbo 等开源模型，下载量破万。	
字节跳动	2024.12 – 2026.12
• 贡献 Seedream 5.0 Pro，主要参与大规模渲染数据与文本编辑相关工作。 • Seedream 5.0 Pro 是面向专业内容生产的多模态图像生成模型，支持高保真视觉合成、指令跟随式图像编辑与基于推理的创作能力。	