

徐晓刚 (Xiaogang Xu)

个人信息

当前工作职位：华为 2012 实验室-中央媒体院-主任工程师

联系电话：+86-17816855719

Email：xiaogangxu00@gmail.com

个人简介：徐晓刚目前在，华为 2012 实验室-中央媒体院，担任主任工程师。此前（2023-1 到 2024-3），徐晓刚在之江实验室担任研究专家（PI），同时兼聘于浙江大学计算机科学与技术学院 CAD 国家重点实验室，担任百人计划研究员，获得之江青年人才托举，之江优秀青年科学基金，浙江省鲲鹏计划，自然科学基金重大项目等支持。2022-7 到 2022-12 期间，徐晓刚在香港中文大学 CSE 系担任 Research Fellow。徐晓刚于 2018 年在浙江大学信息工程专业获得工学学士学位，多次获得国家奖学金，于 2022 年在香港中文大学获得博士学位，师从香港中文大学 IEEE Fellow 贾佳亚教授，获香港政府奖学金。在过去的五年中，徐晓刚在人工智能，计算机视觉领域共发表论文 34 篇（包括 CCF-A 类期刊 4 篇，CCF-A 类会议论文 24 篇），其中 15 篇一作（CCF-A 期刊论文 3 篇，CCF-A 会议论文 10 篇），并且担任多个国际学术会议（包括 CVPR, ICCV, ECCV, NeurIPS, ICLR, ICML, AAAI 等领域顶级会议）和国际学术期刊的审稿人和程序委员会委员。到目前为止申请人谷歌学术（Google Scholar）的引用量为 2000+，GitHub 上论文开源项目累积收获 7000+ star。徐晓刚的研究方向主要是生成式计算摄影，多模态生成和理解大模型，以及安全对齐方向。

教育经历

- 2018.8 - 2022.7 博士：香港中文大学，计算机科学学系
研究方向：计算机视觉(导师：贾佳亚教授，余备教授)
获香港政府奖学金
- 2014.9 - 2018.7 学士：浙江大学，信息与电子工程学院
主修专业：信息工程
- 2015.6 - 2018.7 辅修：浙江大学竺可桢学院
工程教育高级班

工作经历

- 2024.3 - 至今 华为 2012 实验室-中央媒体院-主任工程师
研究方向：AI 赋能的计算摄影，AIGC，多模态 AI
- 2023.1 - 2024.3 之江实验室研究专家&浙江大学百人计划研究员
研究方向：多模态生成和理解

2022.7 - 2022.12 香港中文大学博后研究员&思谋科技研究员
研究方向：多模态恢复

个人技能

Python (精通), C++ (熟悉), latex (精通), AIGC 模型微调 (精通, 包括图像, 3D, 视频模型), 深度学习框架 (Pytorch 精通, Tensorflow 熟悉), JAX (入门), 英语和普通话 (精通)

研究兴趣

本人的主要研究兴趣领域在于人工智能和多模态生成式模型 AIGC。包括

- 1) 多模态数据生成;
- 2) 多模态数据恢复;
- 3) 多模态数据理解;
- 4) 人工智能大模型安全对齐

参与项目

- 1) 之江实验室优秀青年基金项目 (复杂场景下的多模态智能理解, 100 万, 负责人)
- 2) 浙江省鲲鹏计划 (5000 万, 骨干参与成员, 负责模型安全方面)
- 3) 浙江省自然科学基金重大项目 (大模型安全风险护栏理论及关键技术研究, 100 万, 负责人)

发表论文

- **X. Xu**, Y. Wang, L. Wang, B. Yu, and J. Jia. "Conditional Temporal Variational AutoEncoder for Action Video Prediction." IJCV 2023, (CCF-A)
- **X. Xu**, K. Shu, T. Hu, Z. Liu, and H. Bao. "Boosting Image Restoration via Priors from Pre-trained Models." CVPR 2024, (CCF-A)
- **X. Xu**, R. Wang, and J. Lu. "Low-light Image Enhancement via Structure Modeling and Guidance." CVPR 2023, (CCF-A)
- **X. Xu**, R. Wang, C. Fu, and J. Jia. "Deep Parametric 3D Filters for Joint Video Denoising and Illumination Enhancement in Video Super Resolution." AAAI 2023, (CCF-A)
- **X. Xu**, and H. Zhao. "Universal Adaptive Data Augmentation." IJCAI 2023, (CCF-A)
- **X. Xu**, H. Zhao, V. Vineet, S. Lim, and A. Torralba. "MTFormer: Multi-Task Learning via Transformer and Cross-Task Reasoning." ECCV 2022, (CCF-B)
- **X. Xu**, R. Wang, and J. Jia. "SNR-Aware Low-light Image Enhancement." CVPR 2022, (CCF-A)
- **X. Xu**, and N. Xu. "Hierarchical Image Generation via Transformer-based Sequential Patch Selection." AAAI 2022, (CCF-A)
- **X. Xu**, H. Zhao, and Jiaya Jia. "Dynamic divide-and-conquer adversarial training for robust semantic segmentation." ICCV 2021, (CCF-A)

- **X. Xu**, Y. Chen, X. Tao, and J. Jia. “Text-Guided Human Image Manipulation via Image-Text Shared Space.” TPAMI, 2021 (CCF-A)
- **X. Xu**, Y. Chen, and J. Jia. “View Independent Generative Adversarial Network for Novel View Synthesis.” ICCV, 2019, oral (CCF-A)
- **X. Xu**, Y. Zhao, Y. Ding, et al. “No-reference stereoscopic image quality assessment based on saliency-guided binocular feature consolidation.” Electronics Letters 2017
- T. Hu*, **X. Xu***, R. Chu, and J. Jia. “TriVol: Point Cloud Rendering Via Triple Volumes.” CVPR 2023, (CCF-A), * 代表共同一作
- R. Wang*, **X. Xu***, C. Fu, J. Lu, B. Yu, and J. Jia. “Seeing Dynamic Scene in the Dark: A High-Quality Video Dataset with Mechatronic Alignment.” ICCV 2021, (CCF-A), * 代表共同一作
- H. Zhao*, **X. Xu***, Y. Song, et al. “Ranking Users in Social Networks with Higher-Order Structures.” AAAI 2018, (CCF-A), * 代表共同一作
- H. Zhao*, **X. Xu***, Y. Song, et al. “Ranking Users in Social Networks with Motif-based PageRank.” IEEE Transactions on Knowledge and Data Engineering 2019, (CCF-A), * 代表共同一作
- Z. Li, **X. Xu** (通讯作者), S. Lim, and H. Zhao. “UniMODE: Universal Monocular 3D Object Detection.” CVPR 2024, (CCF-A)
- C. Zhang, **X. Xu** (通讯作者), L. Wang, Z. Dai, and J. Yang. “S2WAT: Image Style Transfer via Hierarchical Vision Transformer using Strips Window Attention.” AAAI 2024, (CCF-A)
- L. Yang, **X. Xu**, B. Kang, Y. Shi, and H. Zhao. “Densely Annotated Synthetic Images Make Stronger Semantic Segmentation Models.” NeurIPS 2023, (CCF-A)
- Y. Lao, **X. Xu**, X. Liu, and H. Zhao. “CorresNeRF: Image Correspondence Priors for Neural Radiance Fields.” NeurIPS 2023, (CCF-A)
- H. Wang, **X. Xu**, K. Xu, and R. Lau. “Lighting up NeRF via Unsupervised Decomposition and Enhancement.” ICCV 2023, (CCF-A)
- X. Yang, **X. Xu** (通讯作者), and Y. Chen. “Out-of-domain GAN inversion via Invertibility Decomposition for Photo-Realistic Human Face Manipulation.” ICCV 2023, (CCF-A)
- T. Hu, **X. Xu** (通讯作者), R. Chu, and J. Jia. “Point2Pix: Photo-Realistic Point Cloud Rendering via Neural Radiance Fields.” CVPR 2023, (CCF-A)
- J. Tang, **X. Xu**, S. Hu, and Y. Chen. “High Dynamic Range Image Reconstruction via Deep Explicit Polynomial Curve Estimation.” ECAI 2023, (CCF-B)
- Y. Chen, **X. Xu**, and J. Jia. “Domain Adaptive Image-to-image Translation.” CVPR 2020, (CCF-A)
- Y. Chen, **X. Xu**, and J. Jia. “Homomorphic Latent Space Interpolation for Unpaired Image-to-image Translation.” CVPR 2019, oral, (CCF-A)
- C. Shi, **X. Xu**, S. Ji, K. Bu, J. Chen, R. Beyah, and T. Wang. “Adversarial captchas.” IEEE Transactions on Cybernetics 2021, (CCF-A)
- J. Tang, R. Wu, **X. Xu**, S. Hu, and Y. Chen. “Learning to Remove Wrinkled Transparent Film with Polarized Prior.” CVPR 2024, (CCF-A)
- X. Lai, Z. Tian, **X. Xu**, Y. Chen, S. Liu, H. Zhao, L. Wang, and J. Jia. “DecoupleNet: Decoupled Network for Domain Adaptive Semantic Segmentation.” ECCV 2022, (CCF-B)
- T. Hu, L. Wang, **X. Xu**, S. Liu, and J. Jia. “Self-Supervised 3D Mesh Reconstruction From Single Images.” CVPR 2021, (CCF-A)
- Y. Liu, X. Zhang, and **X. Xu**. “Semantic-Aware Video Color Style Transfer based on Temporal Consistent Sparse Patch Constraint.” ICME 2021, (CCF-B)

- Y. Liu, X. Zhang, and X. Xu. “Reference-based Video Colorization with Multi-scale Semantic Fusion and Temporal Augmentation.” ICIP 2021, (CCF-C)
- L. Yang, B. Kang, Z. Huang, X. Xu, J. Feng, and H. Zhao. “Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data.” CVPR 2024, (CCF-A)
- Y. Liang, X. Yang, J. Lin, H. Li, X. Xu, and Y. Chen. “LucidDreamer: Towards High-Fidelity Text-to-3D Generation via Interval Score Matching.” CVPR 2024, (CCF-A)

荣誉和奖励

- 国家奖学金 (2015)
- ISEE 新人奖学金 (2015)
- 优秀学生称号 (2015, 2016, 2017)
- 三好学生荣誉称号 (2015, 2016, 2017)
- 学业一等奖奖学金 (2015)
- 美国数学建模竞赛二等奖 (2016)
- 浙江大学数学建模竞赛一等奖 (2016)
- 全国大学生数学建模竞赛国家二等奖 (2016)
- 浙江大学电子设计竞赛三等奖 (2016)
- 浙江省物理创新竞赛二等奖 (2015)
- 浙江省高等数学竞赛二等奖 (2015)
- 浙江省政府奖学金 (2016)
- 浙江大学研究与创新二等奖学金 (2016)
- 浙江大学研究与创新一等奖学金 (2017)
- 国家人才培养基地二等奖 (2016)
- 德州仪器助学金 (2017)
- 阿里天池大数据医疗 AI 大赛第 13 名 (2017)
- 浙江大学优秀毕业生 (2018)
- 浙江大学优秀毕业论文 (2018)
- 香港政府奖学金 (2018)
- Large-scale Video Object Segmentation Challenge, ECCV workshop, 第 3 名 (2018)
- 阿里天池安全者挑战赛, 白盒攻击赛道, 第 6 名 (2021)
- 之江实验室优秀青年科研基金 (2023)

教学经历

2018-2019, 秋学期, ENGG1110: Problem Solving by Programming

2018-2019, 春学期, CSCI4190: Introduction to Social Networks

2019-2020, 秋学期, ENGG1110: Problem Solving by Programming

2019-2020, 春学期, CSCI3310: Mobile Computing & Application Development

2020-2021, 秋学期, ENGG1110: Problem Solving by Programming

2020-2021, 春学期, CSCI5109: Image Processing and Computer Vision

实习工作经历

Feb.2020 – Dec.2022 深圳市，思谋科技

研究课题：基于深度学习的视频恢复与增强

- 我主要负责用深度学习技术去完成视频去噪和暗光增强的产品应用。

Sep.2021 - Dec.2021 Microsoft Research，远程实习

研究课题：基于深度模型的多任务学习

- 我们利用基于 transformer 的深度生成模型去完成多任务学习。提升在不同任务上的效果并且减小了网络参数。

Aug.2020 - May.2021 Adobe Research，远程实习

研究课题：多模态图像生成

- 我们设计一个完整的系统，能够用文本场景描述中生成出高分辨率图像。

Dec.2018 - Feb.2020 深圳市，腾讯优图实验室

研究课题：基于深度学习的图像处理技术

- 我设计不同的深度学习算法来完成图像美颜以及图像上色等任务。

June.2018 - Aug.2018 北京市，宽凳科技

研究课题：基于对抗学习的语义分割模型效果提升

- 我们组的目标是利用基于深度学习的语义分割技术来分割城市道路数据，促进自动驾驶技术的发展。
- 我设计不同的算法来产生对抗样本，并且利用对抗训练来提升模型的准确度和鲁棒性。

Apr.2018 – May.2018 浙江省，杭州市，阿里巴巴人工智能实验室

研究课题：基于图像检索技术的识别技术

- 参与阿里天猫精灵项目，主要使得天猫精灵可以自动识别一本童书中的任意一页，以及识别规定厂商的药盒信息。
- 完善后门的图像检索系统，在其中改进传统方法搭建的系统。
- 设计将深度学习方法引入系统的算法。

Sep.2017 - Mar.2018 浙江省，杭州市，商汤科技

研究课题：基于深度学习的视频深度预测

- 利用深度学习的方法训练模型，使得模型可以预测一段视频中任意一帧图像的深度信息，同时达到准确性和实时性的要求。
- 将深度学习方法与传统的 3D 视觉算法结合使用，进一步提升算法效果。

工程研究成果

阿里图灵实验室 AI 智能验证码

- 主要负责利用对抗攻击的技术，完善字符型 AI 智能验证码的设计，使得现有的验证码识别系统无法恶意识别。
- 介绍链接：<https://www.jiqizhixin.com/articles/2018-12-28-7>

思谋“至臻”超高清视频 AI 解决方案

- 主要负责其中面向消费者的桌面视频增强软件 ViQuali 中的超分算法设计，以及老旧视频修复，以及移动端视频画质增强部分。
- 介绍链接：<https://zhuanlan.zhihu.com/p/375433953>

阿里 AI Lab 基于天猫精灵的药品和图书检索系统

- 主要负责设计一套高效的图像检索系统，根据某一个药品或者图书的照片，检索到数据库中相应的信息。

宽凳科技全自动高精度地图

- 主要负责利用基于深度学习的语义分割系统，对于道路场景的车道线等进行语义分割且标识。
- 介绍链接：<https://baijiahao.baidu.com/s?id=1629339024372953804&wfr=spider&for=pc>

腾讯优图实验室老照片自动上色工具

- 主要负责利用基于深度学习的视频上色网络对黑白老照片进行上色。在腾讯优图的多媒体 AI 平台上有过部署，并且在 ICCV2019 的展台上对外展示。