JIALI DUAN

jli.duan@gmail.com ◇ Linkedin ◇ Website

SUMMARY

I'm a senior Research Scientist within Apple's ADM (Apple Diffusion Model) foundation team, specializing in advancing generative models for image and video synthesis through cutting-edge research and innovation. Most recently, we released the **Image Playground** and **Genmoji** features in the Apple Intelligence suite, enabling seamless integration of Generative AI into Apple's Messages, Notes, Emails, and more.

Previously, I was a research scientist at FAIR Labs (Meta), working on scaling foundational models for vision, language and 3D representations, which in turn facilitate generalizations to various downstream tasks. Representative contributions along the line include **Pytorch3D** and **uCO3D** projects.

Prior to that, I was a research scientist at Amazon M5 Multimodal team on building LLMs for e-commerce and customer experiences. My work in this period led to 3 CVPR and 1 NeurIPS **research publications**.

EDUCATION

2017 - 2021 **PhD in Electrical and Computer Engineering** — University of Southern California

2014 - 2017 **MSc in Computer Science** — Chinese Academy of Sciences

2010 - 2014 **BS in Information Engineering** — East China University of Science and Technology

WORK EXPERIENCE

Multimodal Generative Modeling Researcher — 12/2023 - Present
Machine Learning, Apple — *Cupertino, CA*

Research and production in the domain of diffusion and multimodal large-language models. My works include,

- Build multimodal foundational models to facilitate generative AI experiences, such as **Image Playground, Genmoji** that have been shipped in iOS 18, iPadOS 18 and macOS Sequoia.
- Lead the development of **Apple on-device video diffusion model** end-to-end, from general scaling issues to efficiency/optimization aspects, down to corner cases including model bias and safety mitigations.

Senior Researcher — 08/2023 - 12/2023
Graphics, Tencent America — *Palo Alto, CA*

Responsible for doing research on the latest development of generative AI that can help expedite the generation of 3D assets used in gaming. My role is to develop machine learning algorithms and models that facilitate 3D reconstruction given user's inputs, such as a 2D image or a textual description.

- Research on Image-to-3D Reconstruction for generic objects and characters in the game engine.
- Implement reconstruction pipeline based on 2D and 3D diffusion priors for 2D-lifting reconstruction models.

Research Scientist — 06/2022 - 07/2023
FAIR Labs, Meta AI — *Menlo Park, CA*

I work in FAIR Labs of Meta AI, responsible for fundamental research on foundation models in multi-modality and 3D. My role is mainly about developing large language models to push the boundary of AI.

- Member of Onevision dedicated to push the boundary of the next generation of foundational models, capable of visio-linguistic understanding in space/time and 2D/3D.
- Contributed to model-scaling of a 3D counterpart of Segment-Anything model, for learning category-agnostic 3D reconstruction priors. The work spans from SfM stack e.g., hloc, colmap integration, rendering and data curation to NeRF/Gaussian Splat, and text-to-3D generative models which culminated in **uCO3D** project.
- Contributor and maintainer for **Pytorch3D**. Contributed to the release of v0.7.1 and v0.7.2. Examples highly optimized Cuda/C++/python kernels for Marching Cubes, glTF utilities, and Fisheye Camera components.

Applied Scientist II 06/2021 - 06/2022
M5 Core Modeling, Search Science& AI, Amazon *Palo Alto, CA*

Applying AI technologies to make recommendation algorithms more intelligent in understanding users' needs. I develop recommendation machine learning models to deliver product recommendations to the users.

- Built **M5's LLM Pretraining stack** for generic object embeddings that serve a variety of downstream services at Amazon including product search, click-through-rate (CTR) and improving multi-lingual, multi-task metrics.
- Contributed to the LLM infra stack by customizing and diving into open-source frameworks such as DeepSpeed, Hugging Face and Timm.
- Mentored two interns and published 3 CVPR, 1 NeurIPS papers on multi-modal representation learning.

Research Assistant 08/2017 - 05/2021
University of Southern California *Los Angeles, CA*

- Research on interpretable deep neural networks with feed-forward design. Awarded **JVCI 2021 Best Paper**.
- Proposed a pioneering framework that insinuates the idea of adversarial learning with human robot interaction that hits **USC headline** and **IROS 2019 Best Paper Finalist**.
- Recipent of **CCBR 2016 Best Student Paper**.

SELECTED PUBLICATIONS

CVPR 2025. Xingchen Liu, Piyush Tayal, Jianyuan Wang, Jesus Zarzar, Tom Monnier, Konstantinos Tertikas, **Jiali Duan**, Antoine Toisoul, Jason Y. Zhang, Natalia Neverova, Andrea Vedaldi, Roman Shapovalov, David Novotny."UnCommon Objects in 3D".

Apple ML Summit 2024. **Jiali Duan**, Vignesh Jagadeesh."Mitigating Person Hallucination in Diffusion Models under Constraints".

CVPR 2022. **Jiali Duan***, Liqun Chen*, Son Tran, Jinyu Yang, Yi Xu, Zeng Belinda, Trishul Chilimbi. "Multi-modal Alignment using Representation Codebook"

CVPR 2022. Jinyu Yang, **Jiali Duan**, Son Tran, Liqun Chen, Yi Xu, Zeng Belinda, Trishul Chilimbi. "Multi-modal Representation Learning with Triple Contrastive Learning"

ICPR 2022. Xiaoyuan Guo*, **Jiali Duan***, C.-C. Jay Kuo, Judy Gichoya, Imon Banerjee. "Augmenting Vision Language Pretraining by Learning Codebook with Visual Semantics"

ICMR 2022. Xiaoyuan Guo, **Jiali Duan**, Saptarshi Purkayastha, Hari Trivedi, Judy Gichoya, Imon Banejee. "OS-CARS: An Outlier-Sensitive Content-Based Radiography Retrieval System".

CVPR 2021. **Jiali Duan**, Yen-Liang Lin, Son Tran, Larry Davis, C.-C. Jay Kuo. "SLADE: A Self-Training Framework for Distance Metric Learning".

SCMLS 2020. **Jiali Duan**, Xiaoyuan Guo, Son Tran, C.-C. Jay Kuo. "Fashion Compatibility Recommendation via Unsupervised Metric Graph Learning".

IROS 2019. **Jiali Duan***, Qian Wang*, Lerrel Pinto, C.-C. Jay Kuo, Stefanos Nikolaidis. "Robot Learning via Human Adversarial Games".

JVCI 2018. C.-C. Jay Kuo, Min Zhang, Siyang Li, **Jiali Duan**, Yueru Chen. Interpretable Convolutional Neural Networks via Feedforward Design.

ACM-TOMM 2017. **Jiali Duan**, Shuai Zhou, Jun Wan, Xiaoyuan Guo, Stan Z.Li. A Unified Framework for Multi-Modal Isolated Gesture Recognition.

ACCVW 2016. **Jiali Duan**, Jiali Duan, Shengcai Liao, Xiaoyuan Guo, Stan Z. Li. Face Detection by Aggregating Visible Components.

CCBR 2016. **Jiali Duan**, Shengcai Liao, Shuai Zhou, Stan Z. Li. Face Classification, A Specialized Benchmark Study.

PROFESSIONAL SERVICE

- Associate Journal Editor for APSIPA
- Reviewer for CVPR, ECCV, ICCV, ICML, NeurIPS, ACL, EMNLP, TOMM, RA-L, ICIP
- Presenter for USC Robotics Open House 2019

AWARDS

- Best Paper Award in the 2021 Journal of Visual Communication and Image Processing.
- Best Paper Award in the 2019 International Conference on Intelligent Robots and Systems.
- Best Student Paper Award for CCBR 2016.