JIALI DUAN

+1(213) 204-0380 ◇ Mountain View, CA
jli.duan@gmail.com ◇ Linkedin ◇ Website

SUMMARY

I am a researcher with a strong track record in Computer Vision and Artificial Intelligence. I am particularly interested in **multi-modality learning**, **foundation models**, and **generative AI**. I am a quick learner and have a broad interest in new researches and technologies. I have experience in a variety of research projects, including **self-supervised learning**, **vision-language pretraining**, and **human-robot interaction**. I am also proficient in a variety of research tools, including Pytorch3d, Detectron2, Mujoco, Blender, Unity ML, Colmap, Opengl etc.

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

RESEARCH EXPERIENCE

Research Scientist — Jun. 2022 - Present
FAIR Labs, Meta AI — *Menlo Park, CA*

- Contributed to data-scaling and model-scaling efforts of **Text-to-4D** by building a blender-based rendering pipeline and a structure-from-motion (SfM) based reconstruction pipeline. Explored a 3D counterpart of Segment-Anything model for learning category-agnostic 3D reconstruction priors.
- Re-implemented a customized **NeRF** library with reusable research components for 3D volumetric rendering and acceleration.
- Prototyped a **diffusion-inspired** vision-language grounding framework for foundational pretraining.
- Diff reviewer and Contributor for Pytorch3d v0.7.1 and v0.7.2. Examples include Cuda/C++/python kernels for Marching Cubes, glTF utilities, and Fisheye Camera components.
- Lead the 3D data crowd-sourcing efforts by deploying Amazon Mturk websites and building the structure-from-motion stack including hloc, colmap integration, 3D animation and selection criteria metrics.

Applied Scientist II — June. 2021 - Jun. 2022
M5 Core Modeling, Search Science& AI, Amazon — *Palo Alto, CA*

- Performed large multi-modal pretraining 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 1 NeurIPS, 2 CVPR papers on multi-modal representation learning.

Applied Scientist Intern — Summer of 2019, 2020, 2021
A9, Amazon — *Palo Alto, CA*

- Developed one of the earliest semi-supervised metric learning frameworks for leveraging both labeled and unlabeled data in teacher-student distillation that achieved SOTA in retrieval tasks.
- Published 1 CVPR paper on SSL and 1 SCMLS paper on fashion compatibility recommendation with graph.

Research Intern — April. 2017 - Jul. 2017
AuthenMetric — *Beijing, China*

- Developed real-time and time-coherent portrait segmentation algorithms on mobile devices.

Research Assistant Sep. 2017 - May. 2021
University of Southern California *Los Angeles, CA*

- Conducted research on interpretable deep neural networks with feed-forward design that was awarded JVCI 2021 Best Paper Award.
- 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.
- Conducted research and engineering on interactive high-resolution portrait manipulation using generative adversarial networks (GANs).

SELECTED PUBLICATIONS

NeurIPS 2022. Changyou chen, Jianyi Zhang, Yi Xu, Liqun Chen, **Jiali Duan**, Yiran Chen, Son Tran, Belinda Zeng, Trishul Chilimbi."Why do We Need Large Batchsizes in Contrastive Learning? A Gradient-Bias Perspective".

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
- Volunteer for USC Robotics Open House in 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.