

Shan Yang

Research Scientist — Multimodal RL & Reasoning Post-Training

alexyangshan@gmail.com | shanyang.me | [Scholar](#) | [GitHub](#) | [Hugging Face](#)

PROFESSIONAL SUMMARY

- Research scientist focused on **multimodal RL post-training** and visual-language reasoning, with 8+ years across foundation models at Adobe, Amazon, and Google Research
- **Sole-author arXiv 2026** on RL post-training for visual reasoning (**Physics-R1**: GSPO+DAPO, **+18.3 pp** on the held-out olympiad benchmark over Qwen3-VL-8B base); ICML 2026 co-author on test-time reasoning (**+35.8%** Math-500)
- Open-source releases at huggingface.co/shanyangmie: PHYSCORP-A (6.4K audited multimodal physics corpus) and PHYSOLYM-A (500-problem held-out olympiad eval, 99.8% novel-source); prior co-first author at ICCV (AI Choreographer, **700+** citations) and second author at NeurIPS (Attention Bottlenecks, **800+** citations)

TECHNICAL SKILLS

RL post-training: GRPO, GSPO, DAPO, DPO, RLHF, verl, MM-Eureka, reward shaping, LLM-as-judge protocols, audit-aware curricula

Multimodal foundation models: vision-language transformers, autoregressive video / diffusion, test-time training, attention-based fusion

Infra: PyTorch, FSDP / DeepSpeed, vLLM, distributed training at 10B+ params, HuggingFace, Croissant 1.0 + RAI dataset packaging

SELECTED PUBLICATIONS

Physics-R1: An Audited Olympiad Corpus and Recipe for Visual Physics Reasoning

S Yang

arXiv 2026 (2605.14040) — project / code / data

Sole-author. Three-stage corpus audit (Jaccard \rightarrow mxbai \rightarrow Haiku-4.5 judge), 500-problem novel-source held-out olympiad benchmark, and GSPO+DAPO recipe trained at \sim 360 H200-GPU-hours over 3 seeds on Qwen3-VL-8B; +18.3 pp on PhysOlym-A, +15.7 pp on PhysReason. Method follow-up in preparation.

Beyond Test-Time Training: Learning to Reason via Hardware-Efficient Optimal Control

P Wang, S Yang, X Wang, T Xiao, X Liu, C Yu, Y Lou, P Li, Z Wang, MC Lin, R Vidal

ICML 2026

Linear recurrent architecture leveraging optimal control theory for LLM test-time reasoning (+35.8% Math-500, +10% AIME25).

ViLA: Efficient Video-Language Alignment for Video Question Answering

X Wang, J Liang, CK Wang, K Deng, Y Lou, MC Lin, S Yang

ECCV 2024

ICAR: Image-based Complementary Auto Reasoning

X Wang, A Liang, J Liang, M Lin, Y Lou, S Yang

AAAI 2024

Attention Bottlenecks for Multimodal Fusion

A Nagrani, S Yang, A Arnab, A Jansen, C Schmid, C Sun

NeurIPS 2021 — **800+** citations

AI Choreographer: Music Conditioned 3D Dance Generation with AIST++

R Li, S Yang*, DA Ross, A Kanazawa*

ICCV 2021 — **700+** citations

Physics-Inspired Garment Recovery from a Single-View Image

S Yang, Z Pan, T Amert, K Wang, L Yu, T Berg, MC Lin

ACM ToG 2018 — **100+** citations

Modeling Context in Referring Expressions

L Yu, P Poirson, S Yang, A Berg, T Berg

ECCV 2016 — 1600+ citations

RESEARCH IN PROGRESS

Reasoning-method paper (Physics-R1 follow-up): Novel multimodal RL post-training method extending the GSPO+DAPO recipe with judge-robust correctness signals and audit-aware curricula; in preparation, submission Q3 2026.

ENGINEERING & RESEARCH EXPERIENCE

Adobe Foundry

Staff Applied Scientist

San Jose, CA

March 2026 – Present

- Post-training research for Adobe Foundry’s generative AI foundation models powering Firefly: SFT, RLHF / DPO, LoRA, distillation, test-time training
- Owning evaluation protocols and preference-data pipelines for the post-training stack; close collaboration with the foundation-model training team on creator-aligned image / video generation

Amazon (Prime Video & Search)

Senior Applied Scientist Tech Lead

Palo Alto & Sunnyvale, CA

September 2021 – March 2026

- **Prime Video (2025–2026):** Led GenAI Live Action Studio — autoregressive diffusion and training infrastructure for **10B+** parameter video generation; advanced controllable generation via test-time training, teacher-diffusion-forcing, and RL post-training
- **Search (2021–2025):** Pioneered multimodal fusion architecture (video / text / image / structured) for Amazon Video Search, deployed at scale; tech-led ViLA (ECCV 2024, last author) and ICAR (AAAI 2024, last author); led Schema-of-Thought research on multimodal LLM robustness (+28% hallucination reduction)

Google Research

Senior Research Software Engineer

Mountain View, CA

June 2017 – September 2021

- Co-first author at ICCV (AI Choreographer + AIST++ dataset, **700+** citations) and second author at NeurIPS (Attention Bottlenecks, **800+** citations); drove research agenda on music-conditioned 3D motion synthesis and multimodal generative modeling
- Owned 2D/3D pose reconstruction models deployed on GCP at scale

OPEN-SOURCE ARTIFACTS

PhysCorp-A — 6.4K-record three-stage-audited multimodal physics corpus (HF)

PhysOlym-A — 500-problem held-out olympiad benchmark, 99.8% novel-source, EN/ET bilingual subset (HF)

Physics-R1 code — audit pipeline + GSPO+DAPO RL recipe + Sonnet-as-judge eval harness (GitHub)

AIST++ — 1,408-clip motion-music dataset for 3D dance generation (GitHub, Google / ICCV 2021)

EDUCATION

University of North Carolina at Chapel Hill

Ph.D., Computer Science

Chapel Hill, NC

2012 – 2017

Shanghai Jiao Tong University

B.E., Software Engineering

Shanghai, China

2008 – 2012