

Chenwei Xu

☎ (+1)614-254-0561 | ✉ cxu@u.northwestern.edu | 🏠 <https://chenwei-1999.github.io> | 🎓 Google Scholar

EDUCATION

Northwestern University, Ph.D. in Statistics

Evanston, IL, Sep. 2024–Jun. 2029

Advisor: Prof. Han Liu

Northwestern University, M.S. in Computer Science

Evanston, IL, Sep. 2022–Mar. 2024

The Ohio State University, B.S. in Data Analytics

Columbus, OH, Aug. 2018–May 2022

EXPERIENCE

Dolby Laboratories

Research Collaborator

Remote

Aug. 2025–Present

- Led first-author CVPR 2026 sparse-video VQA work, replacing uniform sampling with adaptive evidence retrieval, **summary-state memory**, confidence-based stopping, and compact multimodal reasoning loops across long-video QA benchmarks.
- Raised GPT-4o VideoEspresso accuracy from 26.4% to 48.9% (+22.5pp) using 7.99 frames/video; **RL fine-tuning** lifted NExT-QA from 31.7% to 51.3% while lowering runtime under matched sparse-evidence budgets and comparable inference runs.
- Built **LLM/VLM evaluation** and model diagnostics pipelines across GPT-4o, Qwen, and LLaVA-family models, turning reasoning failures into auditable error cases, benchmark tables, controlled ablations, and safety-oriented reliability reports.
- Curated frame-level evidence traces and answer rationales for long-video QA cases, using localized visual support, summary-state updates, and cross-model disagreement to identify multimodal search and reasoning bottlenecks.

FermiLab

Research Collaborator

Evanston, IL

Mar. 2023–Mar. 2024

- Co-first-authored NeurIPS 2023 ML4Phy work that recast high-frequency beam regulation as a **tracking-control MDP** with PPO optimization, neuralized-PID policy features, and SDF-based control metrics for accelerator regulation.
- Evaluated PID/PPO baselines over nine seeds and 430-step spills, improving SDF by 13.6% over unregulated runs and 1.6% over PID.
- Built reproducible simulation, reward-sweep, and metric-reporting code to separate controller gains from simulator variance, making reinforcement-learning comparisons defensible across repeated seeds, baselines, and fixed evaluation protocols.
- Connected simulator outcomes to interpretable control diagnostics, showing when neuralized-PID structure improved spill uniformity and when plain PPO behavior remained sensitive to reward design and training variance.

Ensemble.AI (Start-up)

Machine Learning Engineer Intern

Evanston, IL

Mar. 2023–Aug. 2023

- Built **AWS** deployment for a start-up Python analytics library using EC2, S3, IAM, Cognito, user-scoped access, and two-hour session timeouts.
- Secured proprietary Python code with least-privilege IAM roles, Cython encryption, Python unittest checks, and reproducible validation.

SELECTED PUBLICATIONS

Towards Sparse Video Understanding and Reasoning. **CVPR**, 2026. [Chenwei Xu](#), Zhen Ye, Shang Wu, Weijian Li, Zihan Wang, Zhuofan Xia, Lie Lu, Pranav Maneriker, Fan Du, Manling Li, Han Liu. [\[Link\]](#)

- Improved long-video VQA under tight frame budgets by designing REVISE with adaptive frame retrieval, **summary-state memory**, and confidence-based stopping, replacing dense sampling with compact evidence traces that preserve multi-step temporal reasoning.
- Raised GPT-4o VideoEspresso by +22.5pp and NExT-QA by +19.6pp with EAGER, an annotation-free **RL fine-tuning** reward, while reusing the same retrieval/stopping interface across proprietary and open-source VLMs for traceable reliability analysis.

VideoCritic: Diagnosing and Localizing Reasoning Errors in Video-Language Models. **CVPR VidLLMs Workshop**, 2026. [Chenwei Xu](#), Jianshu Zhang, Shang Wu, Lie Lu, Pranav Maneriker, Fan Du, Manling Li, Han Liu.

- Converted video QA failures into reproducible diagnostics by building a **VLM auditing** workflow for evidence localization, reasoning-error categorization, cross-model comparison, and safety-oriented reports that expose where reasoning chains break.
- Standardized error analysis by linking model answers to localized visual evidence, enabling targeted ablations, clearer failure-mode reporting, and reusable case records for prompt, model, and dataset-regression comparisons.

BiShop: Bi-Directional Cellular Learning for Tabular Data with Generalized Sparse Modern Hopfield Model. **ICML**, 2024. [Chenwei Xu*](#), Yu-Chao Huang*, Jerry Yao-Chieh Hu*, Weijian Li, Ammar Gilani, Hsi-Sheng Goan, Han Liu. [\[Link\]](#)

- Addressed feature sparsity and non-rotational tabular structure by building bi-directional row/column learning modules with **sparse Hopfield memory**, improving structured-data benchmarks while reducing reliance on heavy HPO.
- Captured intra-feature and inter-feature interactions through adaptive sparse memory and multi-scale representation learning, making Hopfield-style retrieval practical for heterogeneous tabular datasets with categorical, numerical, and missing features.

PhyPrompt: RL-based Prompt Refinement for Physically Plausible Text-to-Video Generation. arXiv, 2026. Shang Wu, [Chenwei Xu](#), Zhuofan Xia, Weijian Li, Lie Lu, Pranav Maneriker, Fan Du, Manling Li, Han Liu. [\[Link\]](#)

- Improved physics-aware text-to-video prompting by contributing **RL-based prompt refinement** and plausibility evaluation, helping the 7B prompt refiner gain +8.6pp joint success on VideoPhy2 through better motion and interaction prompts.
- Supported a dynamic reward curriculum that improved physical commonsense by +11.0pp while preserving semantic adherence, connecting prompt search, multimodal reward design, semantic fidelity, and video-generation failure modes into measurable RL objectives.

OTHER PUBLICATIONS

AdvEvo-MARL: Shaping Internalized Safety through Adversarial Co-Evolution in Multi-Agent Reinforcement Learning. **ICML**, 2026. Zhenyu Pan, Yiting Zhang, Zhuo Liu, Yolo Yunlong Tang, Zeliang Zhang, Haozheng Luo, [Chenwei Xu](#), Yuwei Han, Jianshu Zhang, Dennis Wu, Hong-Yu Chen, Haoran Lu, Haoyang Fang, Manling Li, Chenliang Xu, Philip S. Yu, Han Liu. [\[Link\]](#)

Attention Sinks and Outliers in Attention Residuals. arXiv, 2026. Haozheng Luo, Haoran Dai, Shaoyang Zhang, Xi Chen, Eric Hanchen Jiang, Yijiang Li, Jingyuan Huang, Chenghao Qiu, [Chenwei Xu](#), Zhenyu Pan, Haotian Zhang, Binghui Wang, Yan Chen. [\[Link\]](#)

Phys4D: Fine-Grained Physics-Consistent 4D Modeling from Video Diffusion. **CVPR 4DV Workshop**, 2026. Haoran Lu, Shang Wu, Jianshu Zhang, Maojiang Su, Guo Ye, [Chenwei Xu](#), Lie Lu, Pranav Maneriker, Fan Du, Manling Li, Zhaoran Wang, Han Liu. [\[Link\]](#)

Adaptive Batch Size Schedules for Distributed Training of Language Models with Data and Model Parallelism. **CPAL**, 2025. Tim Tsz-Kit Lau, Weijian Li, [Chenwei Xu](#), Han Liu, Mladen Kolar. [\[Link\]](#)

Communication-Efficient Adaptive Batch Size Strategies for Distributed Local Gradient Methods. arXiv, 2024. Tim Tsz-Kit Lau, Weijian Li, [Chenwei Xu](#), Han Liu, Mladen Kolar. [\[Link\]](#)

Let's Try Again: Eliciting Multi-Turn Reasoning in Language Models via Simplistic Feedback. **ICML AI for Math Workshop**, 2025. Licheng Liu, Zihan Wang, Linjie Li, [Chenwei Xu](#), Yiping Lu, Han Liu, Avirup Sil, Manling Li.

Open-Ended Multi-Modal Relational Reasoning for Video Question Answering. **IEEE RO-MAN**, 2023. Haozheng Luo, Ruiyang Qin, [Chenwei Xu](#), Guo Ye, Zening Luo. [\[Link\]](#)

SMUTF: Schema Matching Using Generative Tags and Hybrid Features. **Information Systems**, 2025. Yu Zhang, Di Mei, Haozheng Luo, [Chenwei Xu](#), Richard Tzong-Han Tsai. [\[Link\]](#)

On Sparse Modern Hopfield Model. **NeurIPS**, 2023. Jerry Yao-Chieh Hu, Donglin Yang, Dennis Wu, [Chenwei Xu](#), Bo-Yu Chen, Han Liu. [\[Link\]](#)

Feature Programming for Multivariate Time Series Prediction. **ICML**, 2023. Alex Reneau, Jerry Yao-Chieh Hu, [Chenwei Xu](#), Weijian Li, Ammar Gilani, Han Liu. [\[Link\]](#)

AlignAb: Pareto-Optimal Energy Alignment for Designing Nature-Like Antibodies. **NeurIPS**, 2025. Yibo Wen, [Chenwei Xu](#), Jerry Yao-Chieh Hu, Kaize Ding, Han Liu. [\[Link\]](#)

Disentangling Beam Losses in the Fermilab Main Injector Enclosure Using Real-Time Edge AI. **JACoW ICALEPCS**, 2023. Kyle Hazelwood, Jeremy Arnold, Mark Austin, Jose Berlioz, Pierrick Hanlet, Jerry Yao-Chieh Hu, Michelle Ibrahim, Jing Jiang, Han Liu, Ashia Livaudais-Lewis, Seda Memik, Jovan Mitrevski, Vladimir Nagaslaev, Aakaash Narayanan, Dennis Nicklaus, Gauri Pradhan, Andrea Saewert, Brian Schupbach, Kiyomi Seiya, Rui Shi, Alexis Shuping, Mattson Thieme, Randy Thurman-Keup, Nhan Tran, [Chenwei Xu](#). [\[Link\]](#)

FPGA Architectures for Distributed ML Systems for Real-Time Beam Loss De-Blending. **JACoW IBIC**, 2023. M. A. Ibrahim, M. R. Austin, J. M. Arnold, J. R. Berlioz, P. Hanlet, K. J. Hazelwood, J. Mitrevski, V. P. Nagaslaev, D. J. Nicklaus, G. Pradhan, A. L. Saewert, B. A. Schupbach, K. Seiya, R. M. Thurman-Keup, N. V. Tran, A. Narayanan, J. Yao-Chieh Hu, [Chenwei Xu](#), J. Jiang, H. Liu, S. Memik, R. Shi, A. M. Shuping, M. Thieme. [\[Link\]](#)

Edge AI for Accelerator Controls (READS): Beam Loss Deblending. **Fast Machine Learning for Science Workshop**, 2023. Jeremy Arnold, Mark Austin, Jose Berlioz, Dave Brace, Pierrick Hanlet, Kyle Hazelwood, Jerry Yao-Chieh Hu, Aisha Ibrahim, Jing Jang, Han Liu, Seda Memik, Jovan Mitrevski, Vladimir Nagaslaev, Aakaash Narayanan, Dennis Nicklaus, Gauri Pradhan, Andrea Saewert, Brian Schupbach, Kiyomi Seiya, Rui Shi, Alexis Maya-Isabelle Shuping, Mattson Thieme, Randy Thurman-Keup, Nhan Tran, [Chenwei Xu](#). [\[Link\]](#)

Using Deep Learning for the Automated Identification of Cone and Rod Photoreceptors from Adaptive Optics Imaging of the Human Retina. **Biomedical Optics Express**, 2022. Mengxi Zhou, Nathan Doble, Stacey S. Choi, Tianyu Jin, [Chenwei Xu](#), Srinivasan Parthasarathy, Rajiv Ramnath. [\[Link\]](#)

GenoArmory: A Unified Evaluation Framework for Adversarial Attacks on Genomic Foundation Models. arXiv, 2025. Haozheng Luo, Chenghao Qiu, Yimin Wang, Shang Wu, Jiahao Yu, Zhenyu Pan, Weian Mao, [Chenwei Xu](#), Haoyang Fang, Hao Xu, Han Liu, Binghui Wang, Yan Chen. [\[Link\]](#)

RESEARCH INTERESTS

My research develops evaluation and modeling systems for **multimodal large language models**, **sparse video understanding**, and **AI for science**. I am especially interested in turning model failures into measurable diagnostics and interventions, including video-language reasoning, physical plausibility for video generation, benchmark construction, tabular learning, and **reinforcement learning** for scientific control.

TECHNICAL SKILLS

Methods: Machine learning, deep learning, **large language models**, **LLM/VLM evaluation**, **multimodal learning**, **multimodal reasoning**, AI safety evaluation, computer vision/NLP, sparse video understanding, benchmark design, **SFT/LoRA**, **RL fine-tuning**, **RAG**

ML/LLM: **PyTorch**, **Hugging Face Transformers**, LLaMA-Factory, Qwen, LLaVA/Video-LLaVA, VeRL, scikit-learn, A100/H100 GPU evaluation, distributed multi-GPU experiments, cached inference

Workflow: Benchmark construction, model diagnostics, frame-level evidence tracing, cross-model disagreement analysis, failure taxonomy, calibration/error analysis, visual evidence reports, and paper-ready ablation summaries

Data/Cloud: **Python**, MySQL/SQL, R, Java, Docker, AWS (EC2/S3/Lambda), Slurm, Linux, Git, cached model outputs, reproducible evaluation pipelines, model-monitoring tables

SERVICE & HONORS

- **Workshop Area Chair:** ICLR 2026 Workshop ES-Reasoning; CVPR 2026 Workshop MMRAgI, supporting reviewer assignment, paper discussion, decision calibration, and research-quality feedback for multimodal/embodied AI submissions.
- **Reviewer:** NIPS 2024; AISTATS 2025; ICML 2025; COLM 2025; NIPS 2025; NIPS 2025 MTI-LLM; NIPS 2025 Math-AI; AISTATS 2026; ICLR 2026; CVPR 2026; COLM 2026; ICML 2026 (Gold Reviewer), covering ML, statistics, multimodal reasoning, math-AI, and LLM papers.
- **Honors:** SAE AI Mini-Challenge Second Prize, 2021; Veeam Software Endowed Scholarship, 2021–2022; ASA DataFest Finalist, 2022, with project work spanning applied ML, statistical modeling, data analysis, and team-based problem solving.