Benlin Liu

liubl@cs.washington.edu — (+1) 310-948-7242 — liubl1217.github.io

Education

University of Washington, Paul G. Allen School of Computer Science & Engineering Seattle, USA
Ph.D. in Computer Science & Engineering Sep 2021 – Present
Affiliation: UW Graphics and Imaging Laboratory (GRAIL)
Advisor: Ranjay Krishna

University of California, Los Angeles Los Angeles, USA
M.S. in Computer Science Sep 2019 – Jun 2021

Tsinghua University Beijing, China
B.Eng. in Electronic Engineering Aug 2014 – Jun 2018

Research Interests

Building multimodal intelligence that can perceive, reason about, and simulate the dynamic visual world. I am particularly interested in complex video understanding, perception-centric reasoning, and training large multimodal models whose "thinking" is grounded in visual perception and long-term memory.

Publications

** indicates equal contribution*

1. **Prioritizing Perception Improves Complex Video Reasoning**
   Benlin Liu, Arka Sadhu, Hyo Jin Kim, Kejie Li, Yifan Wang, Yuning Chai, Ranjay Krishna, Yuliang Li
   Under review

2. **PerceptionComp: A Video Benchmark for Complex Perception-Centric Reasoning**
   Shaoxuan Li, Zhixuan Zhao, Hanze Deng, Zirun Ma, Shulin Tian, Zuyan Liu, Yushi Hu, Haoning Wu, Yuhao Dong*, **Benlin Liu***, Ziwei Liu†, Ranjay Krishna†
   Under review

3. **Structure From Tracking: Distilling Structure-Preserving Motion for Video Generation**
   Yang Fei, George Stoica, Jingyuan Liu, Qifeng Chen, Ranjay Krishna, Xiaojuan Wang, **Benlin Liu**
   Under review

4. **CapNav: Benchmarking Vision Language Models on Capability-conditioned Indoor Navigation**
   Xia Su, Ruiqi Chen, **Benlin Liu**, Jingwei Ma, Zonglin Di, Ranjay Krishna, Jon E. Froehlich
   Under review

5. **Seeking and Updating with Live Visual Knowledge**
   Mingyang Fu, Yuyang Peng, Dongping Chen, Zetong Zhou, **Benlin Liu**, Yao Wan, Zhou Zhao, Philip S. Yu, Ranjay Krishna
   Advances in Neural Information Processing Systems (NeurIPS), 2025

6. **Visual Representations inside the Language Model**
   Benlin Liu, Amita Kamath, Madeleine Grunde-McLaughlin, Winson Han, Ranjay Krishna
   Conference on Language Modeling (COLM), 2025

7. **Coarse Correspondences Boost Spatial-Temporal Reasoning in Multimodal Language Model**
   Benlin Liu, Yuhao Dong, Yiqin Wang, Zixian Ma, Yansong Tang, Luming Tang, Yongming Rao, Wei-Chiu Ma, Ranjay Krishna
   IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025

8. **Interleaved Scene Graph for Interleaved Text-and-Image Generation Assessment**
   Dongping Chen, Ruoxi Chen, Shu Pu, Zhaoyi Liu, Yanru Wu, Caixi Chen, **Benlin Liu**, Yue Huang, Yao Wan, Pan Zhou, Ranjay Krishna
   International Conference on Learning Representations (ICLR), 2025 *(Spotlight)*

9. **GMValuator: Similarity-based Data Valuation for Generative Models**
   Jiaxi Yang, Wenlong Deng, **Benlin Liu**, Yangsibo Huang, James Zou, Xiaoxiao Li
   International Conference on Learning Representations (ICLR), 2025

10. **Efficient Inference of Vision and Language Instruction-Following Models with Elastic Cache**
    Zuyan Liu, **Benlin Liu**, Jiahui Wang, Yuhao Dong, Guangyi Chen, Yongming Rao, Ranjay Krishna, Jiwen Lu
    European Conference on Computer Vision (ECCV), 2024

11. **TIFA: Accurate and Interpretable Text-to-Image Faithfulness Evaluation with Question Answering**
    Yushi Hu, **Benlin Liu**, Jungo Kasai, Yizhong Wang, Mari Ostendorf, Ranjay Krishna, Noah A. Smith
    IEEE/CVF International Conference on Computer Vision (ICCV), 2023

12. **Unleashing Text-to-Image Diffusion Models for Visual Perception**
    Wenliang Zhao*, Yongming Rao*, Zuyan Liu*, **Benlin Liu**, Jie Zhou, Jiwen Lu
    IEEE/CVF International Conference on Computer Vision (ICCV), 2023

13. **DynamicViT: Efficient Vision Transformers with Dynamic Token Sparsification**
    Yongming Rao, Wenliang Zhao, **Benlin Liu**, Jiwen Lu, Jie Zhou, Cho-Jui Hsieh
    Neural Information Processing Systems (NeurIPS), 2021

14. **RandomRooms: Unsupervised Pre-training from Synthetic Shapes and Randomized Layouts for 3D Object Detection**
Benlin Liu*, Yongming Rao*, Yi Wei, Jiwen Lu, Cho-Jui Hsieh, Jie Zhou
IEEE/CVF International Conference on Computer Vision (ICCV), 2021

15. **Robust Object Detection via Instance-Level Temporal Cycle Confusion**
Xin Wang, **Benlin Liu***, Thomas E. Huang*, Fisher Yu, Xiaolong Wang, Joseph E. Gonzalez, Trevor Darrell
IEEE/CVF International Conference on Computer Vision (ICCV), 2021

16. **Multi-ProxyWasserstein Classifier for Image Classification**
Benlin Liu*, Yongming Rao*, Jiwen Lu, Jie Zhou, Cho-Jui Hsieh
AAAI Conference on Artificial Intelligence (AAAI), 2021

17. **MetaDistiller: Network Self-Boosting via Meta-Learned Top-Down Distillation**
Benlin Liu, Yongming Rao, Jiwen Lu, Jie Zhou, Cho-Jui Hsieh
European Conference on Computer Vision (ECCV), 2020

18. **A integrated optical neural network chip based on Mach-Zehnder interferometers**
Xu Zhao, Zhenming Yu, **Benlin Liu**, Yu Li, Hongwei Chen, Minghua Chen
Asian Communication and Photonics (ACP), 2018

19. **Diffusion Models are Few-shot Learners for Dense Vision Tasks**
Benlin Liu, Luming Tang, Yongming Rao, Wei-Chiu Ma, Ranjay Krishna
Manuscript

Research Experience

University of Washington, Paul G. Allen School of Computer Science & Engineering Seattle, USA
Research Assistant to Prof. Ranjay Krishna Sep 2021 – Present

- Research on multimodal large language models, complex video reasoning, and perception-centric benchmarks.
- Train and evaluate models that perform spatial-temporal reasoning, live video querying, and world simulation with generative video and diffusion models.

University of California, Los Angeles, Department of Computer Science Los Angeles, USA
Research Assistant to Prof. Cho-Jui Hsieh Oct 2019 – Jun 2021

- Worked on visual representation learning and efficient deep learning with limited supervision.
- Designed scalable vision models under resource and robustness constraints.

Industry Experience

Meta Summer 2025
Research Scientist Intern

- Research internship on complex video reasoning.

Google DeepMind Jun 2023 – Oct 2023
Student Researcher

- Worked on large-scale vision and multimodal learning problems.

Momenta.ai May 2018 – Jul 2018
Research Intern

- Lane detection and perception for autonomous driving systems.
- Achieved 1st place in Momenta Lane Detection Challenge.

Academic Service

Conference Reviewer:
NeurIPS, CVPR, ICLR, ICCV, ECCV, WACV, AAAI

Honors and Awards

UW Reality Lab–Meta Fellowship
1st place in Momenta Lane Detection Challenge
Tsinghua Academic Excellence Award
1st place in 17th Electronic Design Contest of Tsinghua University
2nd Prize for the 32nd National Undergraduate Physics Olympiad (Top 5%)

Additional Information

Computer Skills: C/C++, Python, Java, Matlab, SQL, Bash, LaTeX, PyTorch, TensorFlow, Keras, Caffe, JAX
Languages: Mandarin Chinese (native), English (proficient)