Yunong Liu

Homepage: yunongliu.com
Twitter/X: YunongLiu1
Email: yunong.liu20@gmail.com

Research Interests

Computer vision and machine learning with focus on 4D scene understanding, visual reasoning, generative models, and self-supervised learning. Specific interests include:

Generative Models: Emergence and control in video generation models, temporal consistency in video generation
Visual Prompting: Investigating emergent behaviors and controllability in generative models through visual prompting
4D Grounding: 4D grounding and spatial-temporal alignment between 3D models and real-world videos
Self-Supervised Learning: Self-supervised representation learning for video understanding

Education

Stanford University, MSc Computer Science (GPA:3.97) Sep. 2023 - Present
- Research Assistant with **four quarters** of RAship.
  · Published one first author paper at NeurIPS 2024, one co-first author paper in submission.
- Anticipated graduation with Distinction in Research
- Selected Courses: *CS148: Introduction to Computer Graphics and Imaging*, *CS348I: Computer Graphics in the Era of AI, CS381: Sensorimotor Learning for Embodied Agents, CS326: Topics in Advanced Robotic Manipulation, etc.*

The University of Edinburgh, BEng Electronics and Computer Science Sep. 2019 - June 2023
- Graduated in joint degree **with Honors**, ranked 2nd.

The University of Texas at Austin, Electrical and Computer Engineering (GPA:3.82) Jan. 2022 - June 2022
- Study Abroad (Completed 18 Credits/Semester)

Publications [3 (co)first-author papers, 2 published, 1 in progress]

Eyzaguirre C.*, **Liu Y.***, Stojanov S., Gaidon A., Niebles J.C., Wu J. (*Alphabetically) *Tracking Emergence in Video Generation Models.* Ongoing ICCV Submission.

Liu Y., Eyzaguirre C., Li M., Khanna S., Niebles J.C., Ravi V., Mishra S., Liu W.*, Wu J.* *IKEA Manuals at Work: 4D Grounding of Assembly Instructions on Internet Videos.* NeurIPS 2024.

Liu Y.*, Kolluri N*, Murthy D. *COVID-19 Misinformation Detection: Machine Learned Solutions to the Infodemic.* JMIR Infodemiology 2022; 2(2):e38756.

Research Experience

Tracking Emergence in Video Generation Models Stanford Vision and Learning Lab
with Prof. Jiajun Wu, Prof. Juan Carlos Niebles, Cristobal Eyzaguirre *Jan 2024 - Present*

Self-Supervised Points Tracking: Developed intervention-based framework using marker prompts to analyze video diffusion models' ability to track points without training. Implemented marker propagation combining EDM-based sampling with classifier-free guidance using blended noise predictions.

Technical Implementation: Employed inversion methods with noise extrapolation exploiting near-linear latent trajectories. Developed adaptive tracking strategies with dynamic search spaces based on motion magnitude. Created comprehensive evaluation framework measuring tracking accuracy (PTA@k), temporal consistency, and occlusion handling.

4D Grounding of Visual Furniture Assembly Instructions Stanford Vision and Learning Lab
with Prof. Jiajun Wu, Prof. Juan Carlos Niebles, Prof. Manling Li, Dr. Weiyu Liu *June 2023 - June 2024*

Led Dataset Development: As the sole student project leader for the year-long project, led the first dataset enabling spatial-temporal alignment of instructional real-world videos with 3D models. Communicated with a team of 30 annotators, ensuring precision through iterative cross-validation.

4D Grounding on Internet Videos: Establish dense correspondences between real-world instructional videos and 3D models. Formulated novel approaches for handling unconstrained internet videos with varying camera parameters and environmental conditions. Developed cross-frame optimization combining PnP-RANSAC with temporal consistency constraints for camera parameter estimation. Designed iterative refinement methods to handle occlusions and maintain consistent part poses across video segments.

COVID-19 Misinformation Detection University of Texas at Austin
with Prof. Dhiraj Murthy *March 2022 - Jan 2023*

Framework: Developed systematic comparison framework across classical models (SVM, LR, BNB) and pre-trained models (BERT, RoBERTa, XLNet) on 7 dataset combinations. Demonstrated classical models excel with combined datasets while pre-trained models perform best. Ensembled crowdsourced annotations (31,441 votes) with model predictions.

Human-AI Integration: Created large-scale crowdsourced validation dataset (31,441 votes from 756 annotators) with rigorous agreement analysis (Krippendorff $\alpha = 0.428$). Developed methods to leverage low-agreement human data (68.5% agreement) to improve model performance.

Industry Experience

Semiconductor Manufacturing Anomaly Detection NXP Semiconductors
Tianjin, China *May 2021 - Aug. 2021*

Adaptive Validation Framework: Created validation methodology integrating domain knowledge with statistical confidence measures. Designed adaptive thresholding mechanism to handle process variations across production lines.

Probabilistic Anomaly Detection: Researched and implemented ensemble detection system combining unsupervised learning methods (Isolation Forest, One-Class SVM) with manufacturing constraints. Developed probabilistic modeling approach for high dimensional multi-variate sensor data with temporal dependencies.

Wind Farm Performance AI Optimization Zealen AI (Startup)
Beijing, China *Aug 2021 - Sep. 2021*

Multi-Modal Integration: Developed framework combining LIDAR measurements with multi-spectral satellite imagery for wind pattern analysis. Implemented automated feature extraction pipeline for processing high-dimensional meteorological time series data.

Hybrid Forecasting Architecture: Designed and validated forecasting system integrating physical models with learned spatial-temporal features. Created validation framework comparing model performance against traditional physical forecasting methods.

Leadership and Community Engagement

Academic Leadership [University of Edinburgh, 2020 - 2023] E&EE Programme Representative advocating curriculum improvements for 200+ students. Teaching demonstrator for foundational courses including Discrete Mathematics, Data Structures, and Computer Systems. Mentored international students through Global Buddies Program.

Volunteer [Sri Lanka Wildlife Conservation Society (SLWCS) , Summer 2018] Led educational initiatives in rural communities while contributing to wildlife conservation research through data analysis.

Honors and Awards

UT Austin Cockrell School of Engineering Fellowship (Declined) [2023] — A $24,000 fellowship for PhD studies.

Turing Scheme Funding [January 2022] — Competitive funding for international academic and leadership exchange.

Leadership in Student Opportunities Award [July 2021] — For leadership and 50+ hours of community service.

1st Year Class Medal [July 2020] — Top performance student in Electronics and Electrical Engineering.

Other Projects

EMo-Mask: Emotional Controllable Motion Generation [Stanford, 2024 Winter] Developed an emotion-aware extension to MoMask for controllable motion generation. Designed EmotionEmbedder module integrating with transformer architecture at different stages (R-transformer, M-transformer, VQ-VAE). Implemented multi-objective training with MSE and contrastive losses for emotion-motion alignment. Created dataset of 320 emotionally expressive motions across 4 emotions and 2 motion types. Conducted systematic ablation studies comparing integration points and analyzing impact on motion quality.

GPU-Accelerated Sparse Tensor Auto-differentiation [Edinburgh, 2023 May] Extended differentiable programming framework for sparse tensors to GPU environments. Developed CUDA implementations focusing on performance optimization and memory efficiency for sparse operations.

Discourse Relation Analysis [University of Edinburgh, 2022 June - 2023 Jan] Worked with Prof Bonnie Webber. Investigated relationships between non-canonical syntax (preposing) and discourse relations using large language models. Helped with the development of mask-filling tasks with BERT to analyze implicit discourse relations, achieving quantitative evidence for syntactic effects on relation prediction.

Open-Source Contribution [Remote, 2022 May-July] Converted a JAX model to ONNX format to improve its running efficiency, and re-implemented and restructured the code to integrate it into the OpenCV code tree.

Just Dance Everything (Best Course Project Award) [UT Austin, 2022 March-June] Developed an accessible dance game that only requires a webcam.

Energy Monitoring and Control System [2020] Developed a system to monitor energy usage, analyze power quality, and control inductors' current.(China Patent No. 7670030)

Technical Skills

Programming Languages: Python, Java, C, CUDA, JavaScript, HTML, Verilog

ML Framework: PyTorch, JAX, TensorFlow, OpenCV, Detectron2

3D/Graphics: Blender, Unity3D, Open3D, OpenGL, PyTorch3D

Engineering: COMSOL, MATLAB/Simulink, Circuit Design, Embedded Systems

Tools: Weights & Biases, AWS/GCP, SLURM, Linux HPC, Git, Docker, Alibabacloud