

Jintang Xue

PhD Candidate in Electrical & Computer Engineering, USC

Los Angeles, CA · +1 (213) 713-6803
jintangx@usc.edu · jintangxue.github.io
Google Scholar · LinkedIn · GitHub

Summary

PhD candidate in ECE at USC (advisor: Prof. C.-C. Jay Kuo), graduating **January 2027**. Research on **multimodal LLMs for 3D scene and object understanding** and on **efficient, interpretable ML** for NLP and medical imaging; first-author papers at NeurIPS 2026 and WACV 2026, co-authored paper at ECCV 2026. Summer 2026 at **Google (YouTube)** on HDR video quality datasets and no-reference quality modeling. Seeking full-time **Research Scientist / Applied Scientist / MLE** roles in multimodal AI, video, and 3D perception.

Education

University of Southern California PhD in Electrical Engineering (minor: Computer Science) Advisor: Prof. C.-C. Jay Kuo	Los Angeles, CA 08/2023 – 01/2027 (expected)
University of Southern California MS in Electrical Engineering (Honors Program)	Los Angeles, CA 08/2021 – 05/2023
Shanghai University BS in Electrical Engineering	Shanghai, China 09/2017 – 06/2021

Industry Experience

Google (YouTube) , Software Engineering PhD Intern	Mountain View, CA 05/2026 – 08/2026
- Built an HDR video quality dataset for YouTube user-generated content (UGC): reproduced state-of-the-art SDR-to-HDR (inverse tone mapping) and super-resolution models to generate HDR/4K variants, pairing SDR sources with their converted HDR versions and HDR sources with tone-mapped SDR and round-trip conversions. - Curated a 3,000-video subset and put it through subjective quality evaluation to obtain ground-truth ratings for model training and validation. - Trained a no-reference HDR variant of UVQ (YouTube's Universal Video Quality model) on the dataset, raising SRCC against subjective ratings from 0.50 to 0.71.	

Research Experience (USC Media Communications Lab)

Multimodal LLMs for 3D Scene and Object Understanding	09/2024 – Present
3D-PLOT-LLM (NeurIPS 2026): made object parts directly addressable by the LLM through learnable region markers and reserved part tokens in the point encoder's token stream, with under 1M new trainable parameters and no segmentation decoder; beat PointLLM, Kestrel, PARIS3D, and SegPoint on every text-output metric of the 3DComPaT-GrIn part-aware grounded description benchmark, up to +3.03 GPT-4o judge score over PointLLM. - Constructed PartVerse-QA, a part-level QA benchmark with 77K training pairs on disjoint object splits, for part-aware 3D LLMs. Descrip3D (WACV 2026): combined multi-view 2D images, 3D object features, and VLM-generated object-level descriptions, injected at two levels (fused into object embeddings and added to the LLM prompt) for grounded QA, captioning, and spatial reasoning; outperformed prior LLM-based and expert models on the primary metric of five 3D scene benchmarks, including +6.0 CIDEr on ScanQA and +2.7 F1@0.5 on Multi3DRefer over Chat-Scene.	
Unsupervised Camouflaged Object Detection	12/2025 – 06/2026
PASR (ECCV 2026, second author): reformulated camouflaged object detection as background-conditioned deviation reasoning in a frozen DINOv2 embedding space, with no training and no pixel-level annotation; outperformed all unsupervised baselines on S-measure, E-measure, and weighted F across four COD benchmarks, using only a 315-image reference library, 50× fewer images than the previous best unsupervised method.	
Clinical AI for Prostate MRI (collaborative project)	08/2022 – Present
Contributed to model design, tuning, and evaluation for ML-based prostate MRI segmentation and cancer detection using interpretable, lightweight models; co-authored 4 journal articles (<i>BJU International</i>, <i>CMIG</i>, <i>Computers in Biology and Medicine</i>, <i>Urologic Clinics</i>) and 7 AUA/EAU conference abstracts (<i>Journal of Urology</i>, <i>European Urology</i>).	

- Proposed **WordFS** (APSIPA TSIP 2024), a weakly supervised feature-selection framework for word-embedding dimension reduction; compressed GloVe, Word2Vec, and FastText embeddings from **300 to 50 dimensions**, with the 50-d GloVe and FastText vectors still beating their original 300-d versions on word similarity; outperformed PCA- and UMAP-based baselines in **under half the runtime** of the PCA-based method.

Lightweight 3D Point Cloud Classification

05/2022 – 08/2023

- Co-developed **Green-PointHop** (APSIPA TSIP 2023, co-first author), a tiny interpretable point cloud classifier for edge deployment, and **S3I-PointHop** (ICASSP 2023), an SO(3)-invariant single-hop variant of PointHop.

Selected Publications (full list on Google Scholar)

-
- [1] **Jintang Xue**, Xinyu Wang, Yixing Wu, Jingwen Chen, C.-C. Jay Kuo, “3D-PLOT-LLM: Part-Level Object Tokens for 3D Large Language Models,” in *NeurIPS*, 2026.
 - [2] Xinyu Wang, **Jintang Xue**, C.-C. Jay Kuo, “PASR: Pattern-Aware Scene-Conditioned Reasoning for Camouflaged Object Detection,” in *ECCV*, 2026.
 - [3] **Jintang Xue**, Ganning Zhao, Jie-En Yao, Hong-En Chen, Yue Hu, Meida Chen, Suyu You, C.-C. Jay Kuo, “Descrip3D: Enhancing Large Language Model-based 3D Scene Understanding with Object-Level Text Descriptions,” in *WACV*, 2026.
 - [4] **Jintang Xue**, Yun-Cheng Wang, Chengwei Wei, C.-C. Jay Kuo, “Word Embedding Dimension Reduction via Weakly-Supervised Feature Selection,” *APSIPA TSIP*, 13(1), e37, 2024.
 - [5] **Jintang Xue**, Yun-Cheng Wang, Chengwei Wei, Xiaofeng Liu, Jonghye Woo, C.-C. Jay Kuo, “Bias and Fairness in Chatbots: An Overview,” *APSIPA TSIP*, 13(2), e102, 2024.
 - [6] Min Zhang*, **Jintang Xue***, Pranav Kadam, Hardik Prajapati, Shan Liu, C.-C. Jay Kuo, “A Tiny Machine Learning Model for Point Cloud Object Classification,” *APSIPA TSIP*, 12(1), e35, 2023. (Green-PointHop; *equal contribution)
 - [7] Pranav Kadam, Hardik Prajapati, Min Zhang, **Jintang Xue**, Shan Liu, C.-C. Jay Kuo, “S3I-PointHop: SO(3)-Invariant PointHop for 3D Point Cloud Classification,” in *ICASSP*, 2023.
 - [8] Yun-Cheng Wang, **Jintang Xue**, Chengwei Wei, C.-C. Jay Kuo, “An Overview on Generative AI at Scale With Edge-Cloud Computing,” *IEEE OJ-COMS*, 4, 2952–2971, 2023.

Skills

ML / LLM: PyTorch, TensorFlow, JAX, Hugging Face Transformers, LoRA, DeepSpeed, scikit-learn, XGBoost**Vision / 3D / Video:** OpenCV, Open3D, PyTorch3D, FFmpeg (libplacebo)**NLP:** NLTK, spaCy**Programming:** Python, C/C++, Java, MATLAB**Tools / Infra:** Git, Linux, LaTeX, Neo4j**Languages:** Mandarin (native), English (fluent)**Awards, Service, and Teaching**

Second Prize, National Undergraduate Electronic Design Contest (TI Cup), China, 2019

Reviewer, IEEE International Conference on Image Processing (ICIP), 2026

Teaching Assistant, USC: EE569 Digital Image Processing (Spring 2025, Spring 2026), EE510 Linear Algebra for Engineering (Fall 2024), EE559 Machine Learning I (Spring 2024)