Peilin Cai

+1 213 994 5318 | peilinca@usc.edu | caipeilin.com

EDUCATION

University of Southern Californ — Los Angeles, CA
Master of Science in Computer Science — *Aug. 2024 – May 2026*

Wuhan University — Wuhan, Hubei
Bachelor of Engineering in Computer Science — *Sept. 2020 – Jun. 2024*

PUBLICATIONS

Secure On-Device Video OOD Detection Without Backpropagation — Oct. 2025
International Conference on Computer Vision (ICCV) 2025 *[PDF] — [arXiv:2503.06166]*
Shawn Li; **Peilin Cai**; Yuxiao Zhou; Zhiyu Ni; Renjie Liang; You Qin; Yi Nian; Zhengzhong Tu; Xiyang Hu; Yue Zhao

A Personalized Conversational Benchmark: Towards Simulating Personalized Conversations — Oct., 2025
NeurIPS Workshop MTI-LLM Spotlight (Top 5%) 2025 *[PDF] — [arXiv:2505.14106]*
Li Li; **Peilin Cai**; Ryan A. Rossi; Franck Dernoncourt; . . . ; Philip S. Yu; Xiyang Hu; Yue Zhao

RESEARCH EXPERIENCE

The Earth Simulator: Building Scalable Street View World Models — Jun. 2025 – Now
Research Assistant · USC Viterbi & GVL Lab · Advisor: Prof. Yue Wang — *Los Angeles, CA*

- Built a street-view world model that turns a handful of web street-view images into long-horizon, action-conditioned exploration videos along user-defined trajectories.
- Trained purely self-supervised on internet imagery and videos without camera poses or external annotations, and scaling to diverse cities using readily available data.
- Unified image-based rendering with learned generative priors to recover scene structure, hallucinate occluded content when needed, and maintain multi-view and temporal consistency.
- Produced geometry-respecting frames with preserved outdoor high-frequency detail and stable extended rollouts; surpassed reconstruction-only and generation-only baselines, enabling interactive exploration, planning look-ahead, and large-scale synthetic experience generation.

LAM: Language Articulated Object Modelers — Feb. 2025 – Sept. 2025
Co-author · USC Viterbi & iLab · Advisor: Prof. Laurent Itti — *Los Angeles, CA*
Submitted to ICLR 2026

- Co-created a pipeline which formulated articulated-object synthesis as a unified code generation task where geometry and articulations are co-designed from scratch.
- Built an agentic pipeline coordinating *Link Designer* (hierarchy reasoning), *Geometry & Articulation Coders* (code writing, compilation, debugging), and *Geometry & Articulation Checkers* (self-correction).
- Used code as an interpretable bridge to select joint types and compute precise placements, ensuring correct inter-link relationships.
- Automatically generated complex, procedurally defined articulated objects from text prompts, demonstrating reliability and scalability.

A Personalized Conversational Benchmark: Towards Simulating Personalized Conversations (PersonaConvBench) — Feb. 2025 – May. 2025
Research Assistant · USC Viterbi & FORTIS Lab· Advisor: Prof. Yue Zhao — *Los Angeles, CA*
Submitted to ICLR 2026

- Co-created **PersonaConvBench**, a large-scale benchmark for personalized reasoning and generation in multi-turn conversations. Specifically, the entire technical implementation, data set acquisition and organization, design and execution of all experiments, and multiple iterations of the method.

- Defined three core tasks—sentence classification, impact regression, and user-centric generation—across 10 diverse Reddit-based domains.
- Built unified prompting/evaluation; observed a +198% relative gain in sentiment classification when using personalized history vs. the best non-conversational baseline.
- Released benchmark, evaluations, and code to spur research on long-horizon, user-adaptive LLMs. arXiv:2505.14106

Secure On-Device Video OOD Detection Without Backpropagation (SecDOOD)
Nov. 2024 – March. 2025

Research Assistant · USC Viterbi & FORTIS Lab · Advisor: Prof. Yue Zhao *Los Angeles, CA*
ICCV 2025 (Poster)
- Co-developed **SecDOOD**, a secure cloud–device collaboration framework enabling on-device OOD detection without device-side backpropagation. Specifically, complete technical implementation, model construction and implementation, and the long period iteration of the model designing.
- Implemented a HyperNetwork-based personalized parameter generator to adapt centrally trained models to device-specific distributions without local fine-tuning.
- Designed dynamic feature sampling & selective encryption to encrypt only the most informative channels, reducing overhead while preserving detection quality.
- Ran extensive experiments across datasets/OOD scenarios; achieved performance comparable to fully fine-tuned baselines. arXiv:2503.06166

PLayTR: Geometric Layout-Driven Structured Plane Reconstruction of Indoor Scenes
Jun. 2023 – Jun. 2024

Research Assistant · Wuhan University · Advisor: Prof. Gui-Song Xia *Wuhan, Hubei*
- Established a prototype with reference to established planar reconstruction methods. Improve the model reconstruction accuracy and segmentation accuracy in the planar reconstruction task by introducing the encoded 3D layout information and Manhattan world cues from the 2D raw image into the cross-attention Transformer.
- Improved the depth prediction method for planar surfaces in the original method from direct top-down global prediction to weighted sum of depth maps with masks after plane-by-plane prediction.
- Designed the data generation module to extract the conforming 3D Layout trilinear segment sets from the dataset accurately and screen out invalid data.
- Reorganized the evaluation structure, introduce a new depth supervision approach, replace the old methods with the more efficient loss function, and design the Hungarian matching for the new trilinear segment group corresponding to the plane group.
- The segmentation results have been improved by 0.7 percentage points over the original model without replacing the backbone, and the reconstruction results show that we have solved the significant problem of poor contact between planes.

TECHNICAL SKILLS

Languages: Python, C/C++, JavaScript, HTML/CSS
Developer Tools: Git, Docker, Google Cloud Platform, VS Code, Visual Studio, PyCharm