

# Jiezhi “Stephen” Yang

[stephenyang.com](http://stephenyang.com) | [stephenyang@fas.harvard.edu](mailto:stephenyang@fas.harvard.edu) | 5104955983 | Cambridge, MA

## EDUCATION

**Harvard University, ME. CSE (Thesis Track), GPA: 4.00/4.00** Aug. 2023 - May 2025  
**Massachusetts Institute of Technology, EECS Cross-registration, GPA: 5.00/5.00** Aug. 2023 - May 2025  
**University of California, Berkeley, B.A. Honors CS, Cognitive Science, Major GPA: 4.00/4.00** Aug. 2019 - May 2023  
**Courses:** DL, ML, CV, LLM, AI, OS, Algorithms, Embedded Intel., Distributed Computing, (Inverse) Graphics, Optimizations, etc.  
**Teaching:** Taught Deep Learning, AI, and Data Structures; led discussion sections, exam reviews/drafting, homework parties, etc.

## INDUSTRY EXPERIENCE

**Machine Learning Engineer Intern, Qualcomm Inc. ([qualcomm.com](http://qualcomm.com))** May 2024 - Aug. 2024  
Built QSCENT DL: pivotal deep learning research for novel on-device smell/gas prediction sensor algorithm using edge AI.

- Designed, implemented, and trained a CNN based Siamese network for smell predictions using PyTorch, achieved 97% accuracy, and proved robustness against unknown, in-the-wild environment-induced signal drifts via Siamese network.
- Devised UNet for electrical signal baseline generation and Physics-informed NN, further reducing performance perturbations.
- Developed end-to-end pipelines for large-scale time-series electrochemical data cleaning, feature extraction, and ablations.

**Software Engineer Intern, Robert Bosch GmbH ([bosch.com](http://bosch.com))** May 2021 - Aug. 2021  
Built pred and viz of voxel-wise traffic visibility, by ray-tracing from traffic LIDAR camera data for infrastructure-based self-driving.

- Implemented prediction algorithm in ROS2 C++, sped up 70% by OpenMP, cache blocking, and OcTree acceleration structure.
- Developed algorithms to model FOVs and reliability variances of LIDAR, Radar, and camera with OctoMap in C++.
- Devised an interactive Rviz GUI for configuring and simulating self-driving sensor setup, enhancing autonomous driving sensor testing with real-time FOV, reliability, and visibility renderings. Adopted teamwide, saving 5 hrs manual efforts per test.

## RESEARCH EXPERIENCE

**Deep Learning Research Assistant, Berkeley Artificial Intelligence Research Lab ([bair.berkeley.edu](http://bair.berkeley.edu))** May 2022 - Sep. 2023  
First-authored CARFF (ECCV 2024 Accepted): A deep learning pipeline to predict future 3D scenes from past visual observations.

- Developed a ViT Pose-Conditional VAE w/ PyTorch to encode view-invariant scene latents for CARLA driving datasets.
- Devised mixture density model & NeRF Decoder w/ GUI to forecast and 3D reconstruct future scenes under uncertainty.
- Halved crashes and doubled efficiency in realistic self-driving simulations, where GAN-based models will mode collapse.

Spearheaded 360Long: Wide-range video reconstructions for VR applications by devising efficient multi-sphere images (MSIs).

- Enabled 3D scene capturing in portable devices, via performing NN on scene radii to cluster and redistribute MSI spheres.
- Interpolated MSIs as a plenoptic tunnel for wider reconstruction. Achieved 12% PSNR increase and faster convergence.

**Computer Vision Research Assistant, MIT CSAIL ([csail.mit.edu](http://csail.mit.edu))** Dec. 2023 - May 2024  
First-authored DreamScenes: Text-conditioned 3D gen via multistep diffusion, LLM, ControlNet, and differentiable rendering.

- Implemented geometry conditioned, masked stable diffusion based pipeline with gaussian splatting for 3D reconstruction.
- Generated photorealistic, finely controlled and textured 3D scenes from text for 3D digital content creation and VR, etc.

**Research Fellow, Harvard Visual Computing & Computational Robotics Group** May 2023 - May 2024  
Led IADP: Addressed vision conditioned robotics domain adaptation via invariant-adaptive compositional transformer diffusions.

- Leveraged anti-causal domain shifts to use orthogonal diffusion models to capture domain invariant and dependent information.
- Implemented and trained to push objects (RL tasks), and demonstrated 3x increased performance under few-shot fine turning.

Spearheaded SportsNeRF: A unified network for visual reconstruction of all human players from monocular sports broadcasting.

- Reconstruct via re-id fine-tuned CLIP-based multi-resolution grid NeRF; estimated skeleton transforms from transposed CNN.
- Achieved single-shot visual reconstruction, 70% faster training with no SSIM reduction. Used 8 GPUs for parallel training.

## TECHNICAL SKILLS

**Languages:** Python, Java, C/C++, SQL, JavaScript, LaTeX, RISC-V Assembly, Markdown, etc.

**Libraries/Tools:** PyTorch, OpenCV, sklearn, NumPy, Pandas, Seaborn, Git, OpenMP, ROS2, Docker, Azure, Amazon EC2, YOLO.

## PUBLICATIONS, PATENTS, AWARDS

**CARFF (first author):** Conditional Auto-encoded Radiance Field for 3D Scene Forecasting **ECCV 2024:** [carff.website](http://carff.website)  
**DreamScenes (first author):** 3D Textured Scene Generation Via Compositional Diffusions **Draft Paper:** [tinyurl.com/dreamscenes](http://tinyurl.com/dreamscenes)  
**ALIA: Diversify Vision Datasets with Automatic Diffusion-based Augmentation** **NeurIPS 2023:** [lisadunlap.github.io/aliam-website](http://lisadunlap.github.io/aliam-website).  
**Qualcomm Intern Patent:** Gas Detection by Miniaturized Electrochemical Device via Deep Learning, scheduled filing Q4/2024.  
**Awards:** Berkeley High Distinction; RISE Lab Summer Research Fellowships, CS Honors Society, Term Honors, Valedictorian, etc.