Shi Chen

+86 166-0210-7753 | +1 (341) 258-8001 | shichen0122@163.com | GitHub: github.com/BillyChern

EDUCATION

Fudan University **Shanghai, CN**

B.Eng. Computer Science and Technology Sept. 2022 – July 2026 (expected)

- GPA: 93/100 (Major GPA: 3.80/4.0); Ranking: 11/91
- Relevant coursework: Data Structure, Principles of Computer System, Operating System, Computer Organization and Architecture, Computer Networks, Algorithms Design and Analysis, Computer Graphics, Machine Learning and Pattern Recognition, Computer Vision, Natural Language Processing, etc.
- Honors and Awards: Second Prize, China Undergraduate Mathematical Contest in Modeling (2023); Third Prize, Chinese Undergraduate Mathematics Competitions (2023); Fudan University Scholarship for Academic Achievement: First with Distinction (2024-2025); Funded by the National Science Foundation of China under its Youth Fund (2024; one of the highest honors for undergraduates in China; about 120 undergraduates nationwide every year).

University of California, Berkeley **Berkeley, CA**

Exchange Student Jan. – Aug. 2025

- GPA:3.91/4.0
- Enrolled courses: Programming Languages and Compilers, Robotic Manipulation and Interaction, Advanced Large Language Model Agents.

ENTREPRENEURIAL EXPERIENCE

Intuition Core Inc. **Berkeley & San Francisco, CA**

Co-founder & CTO May 2025 – Present

- Co-founded Intuition Core Inc., an AI & Robotics startup developing infrastructure solutions to accelerate real-world robot deployment.
- Designed a novel Masked Autoencoders (MAE) based world model pretraining architecture for self-supervised learning on both vision and proprioception data; achieved 15%+ downstream policy training efficiency via scalable utilization of low-quality data through curiosity-based learning.
- Built ultra-low latency (<100ms) teleoperation systems over long distances (Hawaii (Kauai Island) –SF, Shanghai – SF, etc.), featuring intuitive interfaces and control frameworks to support decentralized data collection and real-world deployment.
- Secured $950K in pre-seed funding from leading investors, including Founders Fund and Fen Venture at a $24M post-money valuation cap, led both technical strategy and investor relations.

RESEARCH EXPERIENCE

Semantic Segmentation and Completion for LiDAR Signals **Shanghai, CN**

Project Leader | Supervisor: Prof. Weifeng Ge, Fudan University Sept. 2024 – Present

- Proposed a 3D latent diffusion model with a multi-stage data augmentation pipeline to resolve data size limitations and long-tail distribution in sparse LiDAR datasets, implemented in PyTorch. Benchmarked on SemanticKITTI, the approach effectively addresses class imbalance in sparse training data.
- Implemented a lightweight denoising-based approach for bird's-eye view perception of LiDAR point cloud inputs, achieving 5-10% higher perception accuracy (using mIoU as metric) of sparse LiDAR signal inputs for long-tailed and less-seen categories with no additional computational cost compared to previous state-of-the-art methods.

- Designed a multinomial discrete diffusion model that simultaneously performs semantic segmentation and spatial structure completion on sparse LiDAR signal inputs. The model leverages a discrete diffusion denoising process conditioned on classifier-free guidance (CFG) from the bird's-eye view perception results we get above, achieving 39 mIoU and surpassing the previous state-of-the-art performance of 37.9 mIoU.
- Funded by the National Science Foundation of China under its Youth Fund, currently preparing submission to IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI); actively working on deployment of the algorithm on Unitree G1 humanoid and Go1 robot dog to improve their outdoors navigation.

Mesh & Texture Optimization for 3D Generation **Remote**

Individual Summer Research | Advisor: Prof. Rana Hanocka, University of Chicago June – Aug. 2025

- Extended the Continuous Remeshing pipeline to generate high-quality meshes directly from single-frame reference images, eliminating the need for pre-existing goal geometry through two distinct approaches. Approach 1: Applies MV-Adapter for multi-view generation from the input image, followed by DSINE for surface normal estimation, producing the multi-view normal images required for mesh optimization; Approach 2: Leverages OpenLRM's Transformer-based Large Reconstruction Model to predict a NeRF scene from the input image, then renders required multi-view normal images from the resulting NeRF tri-plane representation.
- Introduced vertex color optimization by designing and integrating a new loss term into the original optimization objective.
- Demonstrated improved mesh quality over the baseline Marching Cube algorithm directly applied to LRM-generated NeRF tri-planes. Quantitative validation via crowdsourced user study on Appen showed >70% preference for our method's output quality.

Flying in the Wild, No GPS **Berkeley, CA**

Team Member | Teammate: Francesco Crivelli, Co-founder & CEO of Intuition Core Inc., UC Berkeley Alumni; Advisor: Prof. Shankar Sastry, University of California, Berkeley Apr. – June 2025

- Built a custom quadrotor under the Agilicious framework with specialized hardware: 3D-printed frames and mounts, Intel RealSense D435i camera, NVIDIA Jetson Orin Nano for onboard computer, TMotor F7 flight controller, and TMotor Velox V2306 V2.0 motors, etc.
- Enabled real-time state estimation using the SVO Pro algorithm on dual infrared cameras and IMU sensor inputs from the D435i for GPS-free navigation.
- Fine-tuned Qwen-2.5 3B VLM on mixed simulator and real-world data for onboard deployment on the Jetson Orin Nano. The VLM processes real-time YOLOv8 object detection outputs to generate navigation directives, guiding our pilot algorithm in waypoint planning and state goal estimation for low-level control execution.
- Developed a customized style-transferred simulation environment with Flightmare and Unity. Increase the realism of the simulated environment in order to narrow the Sim2Real gap between our simulator and the real world.
- Verified the feasibility of onboard VLM deployment for real-time quadrotor navigation, reducing hardware costs by several hundred dollars vs. Agilicious baseline. Documented research findings in the paper "Flying in the Wild, No GPS."

Improving Hand Generation Quality in Diffusion Models **Cambridge, UK**

Individual Summer Research | Advisor: Prof. Marwa Mahmoud, University of Cambridge July – Aug. 2024

- Selected for the Cambridge Summer Research Programme (2024), graduating with First with Distinction.
- Proposed a multimodal diffusion pipeline integrating ControlNet into Stable Diffusion for improved hand-over-face gesture generation. Fine-tuned on specialized datasets of hand-over-face gestures and their corresponding gesture masks. Our approach increased MediaPipe Confidence scores (measuring hand generation quality) from 0.248 to 0.556,

representing a 2.24x improvement over baseline.

- Created a large synthetic dataset of hand-over-face gestures using our generating pipeline, comprising over 170,000 images across 809 types of gestures. Quality validation through crowdsourced evaluation on Appen showed 83.5% of images rated as 'good', with Krippendorff's Alpha of 0.61 indicating fair inter-rater agreement. The synthetic dataset demonstrated measurable impact on downstream tasks, achieving 2-5% accuracy improvements in human gesture detection and gesture sentiment analysis.
- Authored the paper "Hi-Fi: A High-Quality Synthetic Hand-Over-Face Gestures Dataset with Multimodal Diffusion", submitted to the 19th IEEE International Conference on Automatic Face and Gesture Recognition (FG2025).

Reconstructing Streets and Augmenting Autonomous Driving **Shanghai, CN**

Researcher | Collaborator: Dr. Xinrui Wang, Johns Hopkins University Jan. – July 2024

- Developed a diffusion model with Pytorch to generate 3D bounding boxes and bird's-eye view perception data on NuScenes dataset; applied inpainting to scale scene diversity, and created a same-sized synthetic dataset.
- Combined synthetic input for MagicDrive, a 2D driving scene video generation model, to create temporally consistent RGB images and videos. Enhanced consistency and realism by fine-tuning Stable Diffusion with ControlNet, conditioning on 3D bounding boxes and bird's-eye view perception results. Achieved 2-4% accuracy improvement on downstream tasks, including BEV perception and 3D bounding box detection on BEVFusion.
- Used NeRF-based Nerfacto for controllable 3D scene rendering, enabling controllable and editable 3D street scene generation, with potential to improve downstream autonomous driving tasks such as perception, planning, and simulation.

SELECTED COURSE PROJECT

Stateful-Agent: A Cross-Platform LLM-Based Agent with Persistent Memory **Berkeley, CA**

Team Member | Course Instructor: Prof. Dawn Song, University of California, Berkeley Apr. – June 2025

- Developed a stateful LLM agent with persistent memory for academic research management, implementing dual-state architecture using SQLite for structured data (user profiles, research lab collections) and ChromaDB vector database for semantic memory through document embeddings and similarity search.
- Built a Chrome extension to auto-fill graduate school applications, scholarship forms, and fellowship applications by using personal research profile and publications, significantly reducing application time.
- Integrated automated paper discovery and management system that fetches, tracks, and summarizes academic papers from arXiv using LangChain + GPT-4o, maintaining conversation summary memory for context continuity across sessions; implemented cross-platform content sharing to LinkedIn and RedNote as an auxiliary feature for research dissemination.
- Open-sourced the complete system (web UI + terminal), achieving 98% success rate in automated posting tasks and enabling researchers to maintain a long-term research context while streamlining their academic workflow. GitHub repositories: github.com/nsd9696/stateful-agent-chrome-extension, github.com/BillyChern/stateful-agent.

SKILLS

- Programming Languages: C, C++, Python, OCaml, Lean (Theorem Proving), Verilog HDL, Pascal, MATLAB
- Frameworks: PyTorch, PyTorch3d, JAX, OpenCV, Numpy, LangChain
- Robotics & Simulation: ROS / ROS2, MuJoCo, OpenAI Gym
- Modeling & Design: Blender (3D Modeling), Unity
- Scientific Typesetting: LaTeX
- Development Tools: Git, Linux, Docker
- English: TOEFL 106, Speaking 25; GRE 336, Quantitative 170