

Shi (Billy) Chen

+86 166-0210-7753 | +1 (341) 258-8001 | shic@andrew.cmu.edu | shic@intuition.dev
shichen.world | github.com/BillyChern | ORCID 0009-0003-7365-6983

EDUCATION

Carnegie Mellon University

M.S. Robotic Systems Development (MRSD), Robotics Institute

Pittsburgh, PA

Aug. 2026 – 2028 (expected)

Fudan University

B.Eng. Computer Science and Technology

Shanghai, CN

Sept. 2022 – July 2026

- GPA 93/100 (major GPA 3.80/4.0); rank 11/91. Supervised by Prof. Weifeng Ge.
- **Coursework:** Data Structures; Principles of Computer Systems; Operating Systems; Computer Organization and Architecture; Computer Networks; Algorithm Design and Analysis; Computer Graphics; Machine Learning and Pattern Recognition; Computer Vision; NLP.
- **Honors:** National Natural Science Foundation of China Youth Fund (2024; Grant 624B1006 – one of the highest honors for undergraduates in China, ~120 nationwide per year); Fudan Scholarship for Academic Achievement, First with Distinction (2024–2025); Second Prize, China Undergraduate Mathematical Contest in Modeling (2023); Third Prize, Chinese Undergraduate Mathematics Competitions (2023).
- **75th Lindau Nobel Laureate Meeting** (Germany, 28 June – 3 July 2026): one of ~600 young scientists worldwide and one of only 10 undergraduates from China, engaging with 70+ Nobel Laureates.

University of California, Berkeley

Exchange Student

Berkeley, CA

Jan. – Aug. 2025

- GPA 3.91/4.0. Coursework: Programming Languages and Compilers; Robotic Manipulation and Interaction; Advanced Large Language Model Agents.

ENTREPRENEURIAL EXPERIENCE

Intuition Core Inc.

Co-founder & CTO

Berkeley & San Francisco, CA | Shenzhen, CN

May 2025 – Present

- Co-founded an AI & robotics startup building infrastructure to accelerate real-world robot deployment; led technical strategy and investor relations.
- In partnership with **KUKA**, put our *intuition-1* model and ISS infrastructure into production on KUKA arms at Midea's dishwasher factory (Shunde, CN), running in-line quality inspection 24/7.
- Designed a Masked-Autoencoder-based world-model pretraining architecture for self-supervised learning over vision and proprioception, enabling scalable use of low-quality data through curiosity-based learning.
- Built teleoperation systems for decentralized data collection over intercontinental links (Kauai–SF, Shanghai–SF), reaching sub-100 ms *one-way* latency, with operator interfaces and control frameworks around them.
- Secured \$750K in pre-seed funding from Michael Kirsten, a Founders Fund LP investing in a personal capacity.

RESEARCH EXPERIENCE

Generative Semantic Scene Completion

Project Leader | Supervisor: Prof. Weifeng Ge, Fudan University

Shanghai, CN

Sept. 2024 – 2026

- Recast outdoor LiDAR semantic scene completion as a generative problem, using discrete diffusion in three roles: synthesizing training scenes, completing a scene from noise, and refining a frozen model in one step.
- Built **PS³**, an offline paired sparse–dense scene-synthesis pipeline (three cascaded coarse-to-fine multinomial diffusions, a Jensen–Shannon distribution filter, a rare-class object bank, and a virtual HDL-64E ray-tracer) in PyTorch, expanding SemanticKITTI training data $2.67\times$ (19,130 real $\rightarrow$ 51,169 pairs).
- Designed **S²D²**, a structured-source discrete-diffusion operator: the frozen base's argmax replaces the noise prior, so the network learns only the residual to ground truth — no base retraining, no distillation, no test-time adaptation.
- One step lifts a frozen SCPNet base to **38.8% mIoU** on the SemanticKITTI hidden test — to our knowledge the best *causal, single-sweep, single-sample* result on that leaderboard, **+2.1 pp** over the previous best published score under that restriction (SCPNet, 36.7). Four correction steps with an eight-view D_4 ensemble reach 39.2% mIoU / 59.0 completion IoU.
- The same one-step correction adds **+1.3 mIoU** to 2D BEV semantic segmentation (34.8 $\rightarrow$ 36.1, val seq 08).
- Preprint: [arXiv:2608.26737](https://arxiv.org/abs/2608.26737) (under review). Code, checkpoints and the PS³ dataset released: github.com/BillyChern/GSSC-S2D2.

Mesh & Texture Optimization for 3D Generation

Individual Summer Research | Advisor: Prof. Rana Hanocka, University of Chicago

Remote

June – Aug. 2025

- Extended the Continuous Remeshing pipeline to produce high-quality meshes from a single reference image, removing the need for pre-existing goal geometry, via two routes: MV-Adapter multi-view generation with DSINE normal estimation; and OpenLRM's transformer reconstruction model, rendering multi-view normals from the predicted NeRF tri-plane.
- Introduced vertex-color optimization through a new loss term in the original objective.
- Improved mesh quality over marching cubes applied to the same LRM tri-planes; a crowdsourced user study on Appen showed >70% preference for our output.

Flying in the Wild, No GPS

Berkeley, CA

Team Member | *Teammates incl. Francesco Crivelli* | *Advisor: Prof. Shankar Sastry, UC Berkeley* *Apr. – June 2025*

- Built a custom quadrotor on the Agilicious framework: 3D-printed frames, Intel RealSense D435i, NVIDIA Jetson Orin Nano, TMotor F7 controller and Velox V2306 motors.
- Enabled real-time state estimation with SVO Pro over dual infrared cameras and IMU for GPS-free navigation, and built a style-transferred Flightmare/Unity simulator to narrow the Sim2Real gap.
- Deployed a Qwen-2.5 3B VLM on the Jetson Orin Nano that consumes real-time YOLOv8 detections to generate navigation directives for waypoint planning and state-goal estimation; cut hardware cost by several hundred dollars against the Agilicious baseline.

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); graduated with First with Distinction.
- Proposed a multimodal diffusion pipeline integrating ControlNet into Stable Diffusion for hand-over-face gesture generation, raising MediaPipe confidence from 0.248 to 0.556 ($2.24\times$ over baseline).
- Created a synthetic dataset of 170,000+ images across 809 gesture types; crowdsourced Appen validation rated 83.5% of sampled gesture types as “good” (Krippendorff's $\alpha = 0.61$). Authored “Hi-Fi: A High-Quality Synthetic Hand-Over-Face Gestures Dataset with Multimodal Diffusion”.

Reconstructing Streets and Augmenting Autonomous Driving

Shanghai, CN

Researcher | *Collaborator: Xingrui Wang (Ph.D. student), Johns Hopkins University* *Jan. – July 2024*

- Developed a diffusion model in PyTorch generating 3D bounding boxes and BEV perception data on nuScenes; used inpainting to scale scene diversity into a same-sized synthetic dataset.
- Fed synthetic input to MagicDrive for temporally consistent driving video, fine-tuning Stable Diffusion with ControlNet conditioned on 3D boxes and BEV results; reported 2–4% gains on downstream BEV perception and 3D detection with BEVFusion.
- Applied NeRF-based Nerfacto for controllable, editable 3D street-scene rendering.

PATENTS

First inventor on three Chinese invention patents (patentee: Fudan University). Two were granted by the China National Intellectual Property Administration on 1 September 2026; the third is under substantive examination.

- **Granted 1 Sept. 2026**, ZL 2026 1 0956628.1 (CN 122473413 B) — method and apparatus for three-dimensional semantic scene completion based on embodied intelligence.
- **Granted 1 Sept. 2026**, ZL 2026 1 0968991.5 (CN 122493464 B) — LiDAR bird's-eye-view semantic segmentation method and apparatus for embodied intelligence.
- **Under substantive examination**, CNIPA 202610956630.9 (published July 2026) — LiDAR point-cloud data augmentation method and apparatus for embodied intelligence.

SELECTED COURSE PROJECT

Stateful-Agent: A Cross-Platform LLM Agent with Persistent Memory

Berkeley, CA

Team Member | *Instructor: Prof. Dawn Song, UC Berkeley* *Apr. – June 2025*

- Built a stateful LLM agent for research management with a dual-state architecture: SQLite for structured data and ChromaDB for semantic memory over document embeddings.
- Added automated arXiv paper discovery and summarization via LangChain and GPT-4o with conversation-summary memory, plus a Chrome extension that auto-fills graduate and fellowship applications from the user's research profile.

SKILLS

- **Languages:** C, C++, Python, OCaml, Lean, Verilog HDL, Pascal, MATLAB.
- **Frameworks:** PyTorch, PyTorch3D, JAX, OpenCV, NumPy, LangChain, spconv.
- **Robotics & simulation:** ROS/ROS2, MuJoCo, OpenAI Gym, Flightmare.
- **Tools:** Git, Linux, Docker, Blender, Unity, L^AT_EX.
- **Tests:** TOEFL 106 (speaking 25); GRE 336 (quantitative 170).