Qu Delin

+86 18274000661 | delinqu.cs@gmail.com | Portfolio | GitHub | LinkedIn | Shanghai, China

Education

Ph.D. Student
Computer Science and Technology
Fudan University, Shanghai, China
Sep 2022 – Jun 2027

Supervisor: Prof. Xuelong Li (IEEE, AAAI, OSA, SPIE Fellow)
Joint Ph.D. at Shanghai AI Laboratory
Research Interests: Embodied Foundation Model, Spatial Intelligence
NSFC Grant and Tencent Scholarship Research Support
Fudan Top Outstanding Ph.D. Student Award.

Bachelor
Computer Science and Technology
Hunan University, China
Sep 2018 – Jun 2022

Supervisor: Prof. Kenli Li (EVP)
GPA: 3.86 (1/198)
National Scholarships, Finalist of Mathematical Contest in Modeling
*Outstanding Graduate and Excellent Graduate Thesis (**TPAMI**).*

BIO

I am a joint-training CS Ph.D. at Fudan University and Shanghai AI Lab, advised by Prof. Xuelong Li and affiliated with the IPEC@Team. My research focuses on **Embodied Foundation Model and Spatial Intelligence**, with a long-term vision of achieving L2-level Physical Intelligence. I'm excited about the prospect of an "**GPT moment of Embodied AI**", where AI systems learn to interact with the physical world and change the world in a more human-like way.

From 2022, **10+** papers are accpted at Top Conferences (TPAMI CVPR NeurIPS RSS), with **7** first/co-authored CCF-**A** and **4** **Spotlight/Oral** (Top 2.6%). My research "Embodied Foundation Model: Open-World Spatial Generalization" is supported by **National Natural Science Foundation of China** (Top 0.01%) and **Tencent Scholarship** (Top 0.1%). I also awarded **National Scholarships** (Top 0.1%), led **Embodied One Vision** Series Model pre-training, and assisted in Trillion Scaling LLM development.

Intern Experience

Embdoied AI Research Intern
Shanghai AI Laboratory
IPEC@Team
Shanghai, China
Sep 2022 – Jun 2027

Mentor: Prof. Xuelong Li, Dong Wang and Jiangmiao Pang
Focus on Open-world Embodied Foundation Models Research,
Pretraining Group Student Leader, Long Term Joint Ph.D.
Published 10 papers on 3D Vision and Embodied Foundation Model (incl. TPAMI/CVPR/NIPS/RSS).

Embdoied AI Research Intern
AgiBot
Shanghai, China
Mar 2025 – Jun 2025

Mentor: Guanghui Ren and Siyuan Huang
Proposed OpenSS2 - the first open-source hierarchical embodied reasoning model, achieving 24% success rate improvement over π_0 on AgiBot G1 through phased flow matching and value guidance.
"OpenSS2: System 1, System 2 Testing-Time-Computing for High-rate, Dexterous Humanoid Control".

LLM Research Intern
Institute of Artificial Intelligence, @TeleAI
Beijing, China
Mar 2024 – Jun 2024

Mentor: Zihan Wang and Prof. Xuelong Li
Member of Model pretraining Team, Specializing in: Large-scale Discriminator Training (5T tokens processed) and Base Model Optimization, Achieved 15% higher Performance Improvement.
Assisted in TeleChat2 - the first Trillion-Level Large Language Model trained on Huawei 910B.

Publication

Main Conference

Delin Qu, et al., (2025), "SpatialVLA: Exploring Spatial Representations for Visual-Language-Action Model", **Robotics: Science and Systems** (RSS). Spotlight, – 42% Params +78% Perf, Over 30K+ Model Usages! https://doi.org/2501/15830 (Camera Ready).

Main Conference

Delin Qu, et al., (2024), "LiveScene: Language Embedding Interactive Radiance Fields for Physical Scene Rendering and Control", **Conference on Neural Information Processing Systems** (Neurips), OmniSim Bench Towards Embodied Interaction.
https://proceedings.neurips.cc.

Chi Yan*, **Delin Qu*** (Project Leader), et al., (2024), "GS-SLAM: Dense Visual SLAM with 3D Gaussian Splatting", **Conference on Computer Vision and Pattern Recognition** (CVPR), (Spotlight, Top 2.6%, 200+Cite).
https://doi.org/10.1109/CVPR52733.2024.01853.

Delin Qu, et al., (2024), "Implicit Event-RGBD Neural SLAM", **Conference on Computer Vision and Pattern Recognition** (CVPR), (Spotlight, Top 2.6%).
https://doi.org/10.1109/CVPR52733.2024.01852.

Delin Qu, et al., (2023), Towards Nonlinear-Motion-Aware and Occlusion-Robust Rolling Shutter Correction", **Proceedings of the IEEE/CVF International Conference on Computer Vision** (ICCV). 52%+ SOTA Improvement!
https://doi.org/10.1109/ICCV51070.2023.00980.

Bang Yan*, **Delin Qu*** (Coauthor), et al., (2023), "Revisiting Rolling Shutter Bundle Adjustment: Toward Accurate and Fast Solution", **Conference on Computer Vision and Pattern Recognition** (CVPR), (No Degeneracy No Constrain!).
https://doi.org/10.1109/CVPR52729.2023.00471.

Journal Article

Delin Qu, et al., (2023), "Fast Rolling Shutter Correction in the Wild", **IEEE Transactions on Pattern Analysis and Machine Intelligence** (TPAMI), (Single Core CPU Beat 4090 RTX).
https://doi.org/10.1109/TPAMI.2023.3284847.

Ready to Launch

Delin Qu*, **Zhaoqin Chen***, **Qizhi Chen***, et al., (2025), "Interleave Vision-Text-Action Scaling for General Robot Control — A Family of Open Large Embodied Models with Modality All in One", **NIPS Under Review**, (eO Series Foundation Models).

Haoming Song*, **Delin Qu***, et al., (2025), "Hume: Introducing System-2 Thinking in Visual-Language-Action Model", **NIPS Under Review**, (O1-Style Thinking in S1S2).

Qizhi Chen*, **Delin Qu*** (Project Leader), et al., (2025), "FreeGaussian: Guidance-free Controllable 3D Gaussian Splats with Flow Derivatives", **ICCV Under Review**, (Control Any Dynamics with No Guidance).
https://freegaussian.github.io.

Projects

FastUMI Universal Interface
Shanghai AI Laboratory
Sep 2024 – Jun 2025

Kehui Liu, Zhaxizhuoma, **Delin Qu**, et al., (2025), "FastUMI: A Scalable and Hardware-Independent Universal Manipulation Interface with Dataset", A substantial redesign of the Universal Manipulation Interface system enabling rapid deployment and delivering robust performance in real-world data acquisition. https://fastumi.com.

Any4LeRobot
Shanghai AI Laboratory
Mar 2025 – Jun 2025

Qizhi Chen, **Delin Qu** (Project Leader), Haoming Song, et al., (2025), "Any4LeRobot: A tool collection for LeRobot", A curated collection of utilities for LeRobot Projects, including data conversion scripts, preprocessing tools, training workflow helpers and etc.
https://github.com/Tavish9/any4lerobot.

Heterogeneous Agent Systems
Shanghai AI Laboratory
Sep 2023 – Jun 2024

Kehui Liu, Zixin Tang, **Delin Qu**, et al., (2024), "Large Model Heterogeneous Intelligent Agent Systems", International Conference on Intelligent Robots and Systems (IROS), A novel LLM-based task planning framework for collaboration of heterogeneous multi-robot systems including quadrotors, robotic dogs, and robotic arms.
https://link.springer.com/article/10.1007.

Awards & Funding

Scholarship

National Scholarship in 2021 (Top 0.1%)
National Scholarship in 2022 (Top 0.1%)
National Inspirational Scholarship in 2019 (Top 0.05%)
Huawei University Scholarship (Huawei 2022 Future Star) (Top 0.5%)
Tencent Scholarship in 2024 (Top 0.2%)
Fudan Excellence Master Scholarship in 2023 (Top 2.5%)
Fudan Top Outstanding Ph.D. Student Scholarship in 2025. (Top 1%)

Research Funding

National Natural Science Foundation of China (NSFC) (Top 0.01%)
"Learning to Grow: Interactive Morophology for Open-world Robot Control", NSFC Grant *No. 624B2044*.
Developed growth-oriented embodied AI that learns interactive morphology through multimodal large models, enabling autonomous skill evolution in open-world environments. Proposed novel training paradigms combining physical simulation with human-like learning mechanisms to overcome data dependency in traditional approaches.

Competition

Finalist Prize of Mathematical Contest in Modeling (Top 1%)
Second Prize in National Internet of Things Design Contest (Top 1%)

Teaching

Teaching Assistant
Object Oriented Programming
COMP130135.04, Fudan University
Mar 2023 – Jul 2023

Boosted coding efficiency by 30% for 100+ students through structured debugging guidance, while designing core algorithmic assignments and leading lab sessions. Earned exceptional teaching evaluation for "*Outstanding Pedagogical Support and Technical Expertise*" from faculty.

Media Presence

Invited Talk
Institute of Artificial Intelligence, TeleAI, *Mar 2025*

Hosted by Prof. Chenjia Bai
SpatialVLA: A spatial-enhanced vision-language-action model that is trained on 1.1 Million real robot episodes. Paradigms and Challenges Towards Generalist Agents System.
delinqu.github.io/talk/202503-teleai/spatialvla-slides

Skills

Language

Chinese: Native, **Hsiang:** Native
English: Fluent

Research Coding

Programming: Python, MATLAB, C/C++, Java, CUDA, NeoVim
Framework: Transformers, Acclerate, LeRobot, Diffusers, NeRFStudio, TRL, DeepSpeed

Academic Service

Conference Reviewer
Sep 2022 – Jun 2025

CVPR 2022–2025, ICCV 2022–2025, ECCV 2023–2025, NeurIPS 2023–2025, ICML 2024–2025, ICLR 2024–2025, TPAMI 2025, AAAI 2024–2025.