Delin Qu

182-7400-0661 | dlqu22@m.fudan.edu.cn | Shanghai

Delin Qu is a second-year Ph.D. student at the School of Computer Science, Fudan University, and a Research Intern at the Shanghai AI Laboratory, advised by Professor Xuelong Li. His research focuses on Perception and Interaction Foundation Model for Embodied Intelligence, including vision language action foundation models, interactive physical world modeling, and simultaneous localization and mapping based on neural rendering. Previously, he authored multiple papers in top conferences and journals such as TPAMI, CVPR, and ICCV. Additionally, he played a key role in several national projects, including Embodied Large Language Model Pretraining and Large Language Model for Distributed Multi-agents, and received numerous honors, including CVPR Highlight and Tencent Scholarship.

EDUCATION

Fudan University   2022.09 - 2027.06
Ph.D., computer science and technology GPA: 3.74/4.0 Outstanding Student   *Shanghai, China*

Hunan University   2018.09 - 2022.06
Bachelor, computer science and technology GPA: 1/208 National Scholarship   *Changsha, China*
and Outstanding Graduate

AWARDS & HONOURS

2022.09 — Now

- Tencent Scholarship in 2023, Fudan University Master's Excellence Scholarship in 2022, Outstanding Student Award in 2023, Fudan University's Outstanding Youth League Member in 2024, Six patents pending, Featured as an outstanding Ph.D. student representative in Shanghai AI Lab's official social media account

2018.09 — 2022.06

- National Scholarship in 2021, National Scholarship in 2020, National Inspirational Scholarship in 2019, Finalist Prize of Mathematical Contest in Modeling, Second Prize of Asia-Pacific Mathematical Contest in Modeling, Second Prize in National Internet of Things Design Contest, Second Prize in Internet Competition of Hunan Province, Excellence Award in the Huawei AI Cloud Cup, Huawei College Scholarship, Huawei Smart Base Future Star, Excellent Graduation Thesis

RESEARCH

six top conference papers as first author (2 Highlights) during 1.5 years

- **GS-SLAM: Dense Visual SLAM with 3D Gaussian Splatting** , Chi Yan*, **Delin Qu***, Dong Wang, Dan Xu, Bin Zhao, Xuelong Li, *CVPR 2024 Highlight*
- **Implicit Event-RGBD Neural SLAM**, **Delin Qu**, Chi Yan, Dong Wang, Jie Yin, Qizhi Chen, Yiting Zhang, Dan Xu, Bin Zhao, Xuelong Li, *CVPR 2024 Highlight*
- **Fast Rolling Shutter Correction in the Wild** , **Delin Qu**, Bangyan Liao, Omar Ait-Aider, Yizhen Lao, **TPAMI 2023**
- **Revisiting Rolling Shutter Bundle Adjustment: Toward Accurate and Fast Solution** , Bangyan Liao*, **Delin Qu***, Yifei Xue, Huiqing Zhang, Yizhen Lao, **CVPR 2023**
- **Towards Nonlinear-Motion-Aware and Occlusion-Robust Rolling Shutter Correction** , **Delin Qu**, Yizhen Lao, Dong Wang, Zhigang Wang, Bin Zhao, Xuelong Li, **ICCV 2023**
- **Ego-motion Estimation for Vehicles with a Rolling Shutter Camera** , Yongcong Zhang, Bangyan Liao, **Delin Qu**, Junde Wu, Lu, Li Xinyu, Wanting, Yifei Xue, Yizhen Lao†, **IEEE Transactions on Intelligent Vehicles 2024**
- **LiveScene: Language Embedding Interactive Radiance Fields for Physical Scene Rendering and Control** , **Delin Qu**﹡, Qizhi Chen﹡, Pingrui Zhang, Xianqiang Gao, Bin Zhao, Dong Wang†, Xuelong Li, **Neurips 2024** (under review)
- **Learning 2D Invariant Affordance Knowledge for 3D Affordance Grounding,** Xianqiang Gao, Pingrui Zhang, **Delin Qu**, Dong Wang, Zhigang Wang, Yan Ding, Bin Zhao, **AAAI 2024** (under review)

PROFESSIONAL EXPERIENCE

Embodied Large Language Model Pretraining   2024.03 - Present
Served as the training team lead for large-scale multimodal model training and   *Shanghai*
optimization, Shanghai AI Laboratory

- Released the InternVLA multi-modal model system for vision-language-action, achieving top two rankings among open-source models in multiple benchmark tests including SimplerEnv, surpassing Google's RT1 & RT2 series and

the latest OpenVLA.

- Designed a portable Gripper operation data collection solution, gathering over 30,000 distributed robot interaction manipulation datasets from a first-person perspective across 600+ scenarios. This work provides further data support for advancing towards the era of general artificial intelligence, with related software, hardware configurations, and data soon to be open-sourced.

Edge-Device Large Language Model Competition, NeurIPS 2024 Challenge　　　　2024.08 - 2024.11
Model architecture design and training, team leader　　　　*Vancouver, Canada*

- Led the team in designing the CrazyLM architecture from scratch, efficiently adapting it for 12GB RAM devices. Conducted training and optimization of CrazyLM on the C4 and Alpaca datasets. As a result, CrazyLM achieved the top three rankings in the global leaderboard of NeurIPS 2024 across multiple evaluation metrics, demonstrating our ability to build and rapidly train large language models under limited resource

Large Language Model for Distributed Multi-agents　　　　2023.01 - Present
Action detection and understanding leading, Algorithm Team Member , Shanghai　　　　*Shanghai*
AI Laboratory

- Proposed the world's first LLM-driven collaborative robot-controlling framework using large models to orchestrate autonomous agent cooperation, achieving a comprehensive mobilization of robots

Humanoid Robot Agent System　　　　2024.03 - Present
SLAM Algorithm Deployment and Dataset Collection, Algorithm Team Member ,　　　　*Shanghai*
Shanghai AI Laboratory

- Leveraging a vast dataset of humanoid interactions, we built an autonomous system featuring humanoid robots, robot dogs, drones, and navigation robots, pioneering a new approach to human-centric embodied intelligence

Optics-driven Drone for unlimited-time flying　　　　2022.08 - 2023.10
ultra-real-time target tracking algorithms based on event cameras, serving as a　　　　*Shanghai*
key member of the algorithm team, Shanghai AI Laboratory

- Leveraged intelligent information transmission processing and high-energy laser technologies to achieve all-weather intelligent visual tracking and autonomous long-range energy replenishment for drones. This advancement enabled autonomous power supply, providing drones with unlimited endurance capabilities

COMP130135.04 Object Oriented Programming　　　　2023.03 - 2023.07
Teaching Assistant, Fudan University　　　　*Shanghai*

- Aiding the professor in teaching, designing algorithmic programs, and conducting courses

SKILLS & CERTICATES

- Programming Languages: Python, **CUDA**, Torch, Jax, C, C++, Java, Lua
- Training Frameworks: Transformers, Diffusers, LeRobot, DeepSpeed, NeRFStudio, ORBSLAM
- Interests: Marathon, Middle/Long-distance Running, Photography, **Neovim**