# Qilin Huang

Shenzhen, China

(+86) 189-0228-7420 | huangqilin2022@mail.sustech.edu.cn | Personal Website


## EDUCATION

**Southern University of Science and Technology (SUSTech)** Shenzhen, China

*Bachelor of Science in Computer Science* **Sep. 2022 – Jun. 2026 (Expected)**

**GPA:** 3.89/4.00 (Rank: 14/162)

**Relevant Coursework:** Artificial Intelligence (A), Natural Language Processing (A+), Probability and Statistics (A), Discrete Mathematics (A+), Computer Organization (A), Computer Networks (A+)

**National University of Singapore (NUS)** Singapore

*Summer Workshop in Computer Vision* **May 2024 – Jul. 2024**

**Grade:** A+

## RESEARCH INTERESTS

- **Structured 3D Visual Reasoning:** Building interpretable models for complex 3D reasoning tasks.
- **Physics-Informed 3D AI:** Learning physical properties from visual data to enable realistic simulation for embodied AI.
- **Multimodal Learning:** Fusing vision and language for robust 3D foundation models.

## RESEARCH EXPERIENCE

**Learning Variational Physical Representations from Visual Features** | Visiting Researcher Philadelphia, PA, USA

GRASP Lab, University of Pennsylvania **Jul. 2025 – Present**

Supervisor: Prof. Lingjie Liu

- Upcoming research on more variable and expressive latent representations of physical properties, moving beyond deterministic point estimates to better capture material uncertainty inherent in visual data.

**Neural Module Network for 3D-VQA** | Project Leader Shenzhen, China

Southern University of Science and Technology **Feb. 2025 – Present**

Supervisor: Prof. Feng Zheng

- Spearheaded research into Neural Module Networks (NMN) for 3D Visual Question Answering (VQA), tackling ScanQA's limited scene diversity to bolster model generalization across unseen environments.
- Developed a tailored NMN framework for 3D scenarios and implemented LLM-driven prompting strategies to convert natural language questions from the ScanQA dataset into executable function programs, a crucial step for aligning inputs with the NMN framework's requirements.
- Currently developing specialized 3D NMN function modules aimed at augmenting model interpretability, structured reasoning, and optimizing overall model performance.

**Hierarchical Concentration Narrowing for 3D-VQA** | Research Assistant Shenzhen, China

Southern University of Science and Technology **Jun. 2024 – Sep. 2024**

Supervisor: Prof. Feng Zheng

- Established the ScanQA environment and executed preliminary experiments, involving the training and assessment of the ScanQA model, to define a performance benchmark for 3D Visual Question Answering (VQA).
- Modified and integrated the Hierarchical Concentration Narrowing (HCN) module into the BridgeQA model architecture, demonstrating its potential to enhance spatial reasoning by suppressing shortcuts via hierarchical supervision, thereby providing valuable groundwork for subsequent development and evaluation.
- Conducted comparative experiments validating HCNQA's effectiveness in enhancing reasoning pathways. HCNQA surpassed the optimal baseline (3D-VisTA) on all metrics, with 4.12% higher EM@1 and a 3.09% CIDEr boost.

*Supported by National Undergraduate Training Program for Innovation and Entrepreneurship.*

**Enhancing Spatial Reasoning in 3D Scene Understanding with LLMs** | Research Assistant Shenzhen, China

Southern University of Science and Technology **Jan. 2024 – May 2024**

Supervisor: Prof. Feng Zheng

- Investigated the inherent limitations of Large Language Models (LLMs) in 3D generalist applications, particularly noting the loss of spatial information resulting from multimodal fusion.
- Proposed integrating textual spatial embeddings derived from a specialized 3D grounding model (EDA) into LLMs to enhance spatial reasoning capabilities.
- Designed model architecture, configured EDA, and modified code to generate spatial embeddings for downstream tasks.
- Deployed and configured LL3DA and LEO model environments, modifying LEO's codebase to process supplementary spatial data for enhanced reasoning performance.
- Identified key challenges in direct multimodal fusion, which directly motivated the subsequent development of the structured, module-based approach in the HCNQA project.

## PROJECTS

**Multi-User Video Conferencing System** | Core Member — Shenzhen, China
Southern University of Science and Technology — **Jan. 2025**

- Developed a real-time video conferencing system in Python using Flask and SocketIO, supporting Client-Server (C/S) and Peer-to-Peer (P2P) frameworks with dynamic mode switching for multiple concurrent users.
- Implemented video and text transmission over TCP/UDP with STUN-based NAT traversal, ensuring low-latency streaming and content isolation across multiple parallel meetings.

**Information Exposure Maximization in Social Networks** | Developer — Shenzhen, China
Southern University of Science and Technology — **Nov. 2024**

- Developed greedy best-first search with incremental updates and simulated annealing algorithms to maximize balanced information exposure in social networks, reducing echo chambers using an efficient incremental evaluation strategy.
- Validated effectiveness through Monte Carlo simulations and benchmarking across diverse graph structures, ensuring robust mitigation of information polarization under budget constraints.

**Masked-Unmasked Face Recognition based on Chain of Thoughts** | Core Member — Singapore
National University of Singapore (NUS) — **May 2024 – Jul. 2024**

- Contributed to the model architecture design, focusing on adaptive local feature selection to improve recognition accuracy under varying conditions.
- Conducted dataset analysis on Georgia Tech Face and CelebA, evaluating feature distinguishability and model performance on masked/unmasked data.
- Developed data preprocessing pipeline, implementing face alignment and feature extraction for mask detection.
- Won the Third Prize in NUS Visual Computing course for outstanding project performance.

**Optimized Matrix Multiplication in C** | Developer — Shenzhen, China
Southern University of Science and Technology — **May 2024**

- Implemented and optimized floating-point matrix multiplication in C, achieving up to 10x performance improvement over naive implementation using techniques like OpenMP, blocking, SIMD (AVX2), and write caching.
- Conducted performance benchmarking with Google Benchmark and compared results against OpenBLAS, surpassing its efficiency for specific matrix sizes; validated functionality across various dimensions with Python-based data analysis.

## HONORS AND AWARDS

**SUSTech Third Class Scholarship** — **Sep. 2023, Sep. 2024**
**Third Prize, Visual Computing Course Project**, National University of Singapore — **Jul. 2024**

## SKILLS AND INTERESTS

**Programming:** C/C++, Python, Java, SQL, JavaScript, Assembly, Verilog
**Tools & Frameworks:** PyTorch, Scikit-Learn, LaTeX
**Languages:** English (TOEFL: 106), Mandarin (Native), Cantonese (Native)
**Interests:** Photography, Anthropology, Running, Volunteering (70+ hours)