Zhangyang Qi (Alex Chi)

Wechat: openainvidia | **Email:** qizhangyang2000@gmail.com | Google Scholar | Homepage | Github

[Education]

The University of Hong Kong (HKU),   PhD in Computer Science   Sep. 2022 - Jul. 2026(Exp)
- **Advisor**: Hengshuang Zhao and Yizhou Yu
- **Research Direction:** `Video Multimodal Large Model,  3D Scene Understanding,  Embodied AI`

Harbin Institute of Technology (HIT),   Bachelor in Information Engineering   Sep. 2018 - Jul. 2022
- **GPA**: 95.53/100 or 3.9706/4.0,   **Rank**: 1/318

[Selected Publications]

GPT4point: A Unified Framework for Point-Language Understanding and Generation

Z. Qi, Y. Fang, Z. Sun, X. Wu, T. Wu, J. Wang, D. Lin, H. Zhao   [CVPR 2024 Highlight]

GPT4Point++:   Advancing Unified Point-Language Understanding and Generation

Z. Qi, Y. Fang, Z. Sun, X. Wu, T. Wu, J. Wang, D. Lin, H. Zhao   [TPAMI 2025 (Minor Revision)]

OCBEV: Object-Centric BEV Transformer for Multi-View 3D Object Detection

Z. Qi, J. Wang, X. Wu, H. Zhao   [3DV 2024]

GPT4Scene: Understand 3D Scenes from Videos with Vision-Language Models

Z. Qi, Z. Zhang, Y. Fang, J. Wang, H. Zhao   [Arxiv 2025]

Tailor3D: Customized 3D Assets Editing and Generation with Dual-Side Images

Z. Qi, Y. Yang, M. Zhang, L. Xing, X. Wu, T. Wu, D. Lin, X. Liu, J. Wang, H. Zhao   [Arxiv 2024]

[Experience]

Shanghai AI Laboratory,   Research Intern – Shanghai, China   Jul 2022 – Until Now
- Research on **3D & video language models**, developing the *GPT4Point*, *GPT4Point++*, and *GPT4Scene*.
- Curated training data for *InternLM-XComposer* series and *V3Det* dataset.

Tencent PCG,   Research Intern – Shenzhen, China   Dec 2021 – May 2022
- Built CLIP-based cross-modal alignment via contrastive learning for image-text matching.
- Designed joint training paradigms enhancing embedding alignment in multimodal retrieval.

[Projects]

Pointcept (1.9K Stars):   A popular Unified 3D perception toolkit.   github.com/Pointcept/Pointcept

- **Point Transformer Series**: From PTv1 to PTv3, continuous optimization has made it a benchmark model in point cloud perception, published at top conferences such as CVPR and NeurIPS.

- **Broad Compatibility**: Supports diverse backbones (e.g., MinkUNet, SPVCNN, Swin3D) and tasks (e.g., segmentation, pretraining), with preprocessed datasets and code.

- **My main Responsibility**: object point cloud understanding, as well as the improvement of indoor scene datasets (ScanNet, ScanNet++) and algorithms.

[Awards]

- Hong Kong PhD Fellowship Scheme (HKPFS) — 2022
- HKU Presidential Scholarship (HKUPS) — 2022

- Top Ten Students of Harbin Institute of Technology — 2021
- Top Ten students of Learning Stars of HIT — 2021
- Top Ten Students of Honor School in HIT — 2022
- Excellent graduate of HIT — 2022

- China National Scholarship — 2020
- The Third Prize of the National College English Contest — 2020
- First Prize of Mathematical Contest in Modeling — 2019
- First prize of the 11th National College Student Mathematics Competition — 2019

Technologies & Service

- **Languages:** Python, PyTorch, Java, JavaScript

- **Academic Services:**

  - Conference Reviewer: CVPR (2024, 2025), ICCV 2025

- **Teaching Experience:**

  - *DASC7606: Deep Learning* (Graduate course) — 2023, 2024 Spring, 2024 Fall

- **Foreign Languages:**

  - TOEFL iBT: 110   IELTS: 7.5   GRE: 330+4.0
  - Cantonese: Conversational proficiency