

XINGTONG GE

Ph.D. student at HKUST since Fall 2025

Homepage · Google Scholar · GitHub · xingtong.ge@gmail.com · Birth: 2001.09

RESEARCH INTERESTS

Diffusion/Flow Distillation: efficient image/video generation, distribution matching, cache-aware training

Post-training & Alignment: preference optimization for diffusion/flow and autoregressive video models

Neural Representation & Compression: image/video compression, Gaussian splatting, task-aware coding

EDUCATION

Hong Kong University of Science and Technology (HKUST)

Aug 2025 – Jun 2028 (*expected*)

Ph.D. in Electronic & Computer Engineering

Advisor: Prof. Jun Zhang (IEEE Fellow)

Beijing Institute of Technology (BIT)

Sep 2022 – Jun 2024

M.S. in Software Engineering

Beijing Institute of Technology (BIT)

Sep 2018 – Jun 2022

B.S. in Computer Science & Technology

Rank: Top 10%

Outstanding Graduation Thesis Award

Top 5%

EXPERIENCE

Vivix AI

Aug 2025 – Present

Research Intern

Shenzhen/Hong Kong, China

- Worked on distillation algorithms and streaming generation training for Vivix video models; validated self-consistent distribution matching and KV-cache-aware training on Vivix-Turbo, integrated it into the business-model training pipeline, and extended it to Wan 2.1 and Self-Forcing as Salt (ECCV 2026).
- Validated multi-stage adversarial training for few-step video generation, including attention-based discriminator heads; advanced an LTX-2 autoregressive post-training pipeline with coarse-to-fine optimization over training, step distillation, and long-video refinement, outperforming Causal Forcing++ in internal evaluations.
- Co-authored AR-CoPO (ECCV 2026) and Neighbor GRPO (CVPR 2026), designing contrastive policy optimization objectives for autoregressive video generation and flow models to improve preference-alignment stability and convergence.

SenseTime Research

Feb 2024 – Aug 2025

Algorithm Researcher (promoted from Research Intern), Base Model Group

Beijing, China

- Deeply involved in distillation-acceleration R&D for SenseTime's *Miaohua* text-to-image models; built segment-wise distillation and adversarial training pipelines, launched *Miaohua 1.1/1.1 Anime/1.2/2.1*, and achieved over 7× generation speedup while maintaining high visual quality.
- Proposed SenseFlow, a scalable distribution-matching distillation framework for large text-to-image models; addressed DMD limitations on large models, achieved leading one- to four-step generation and multi-metric SOTA on FLUX, SD3.5 Large, and benchmarks; accepted to ICLR 2026 and improved through GRPO-style post-training.
- Co-developed an on-device text-to-image system for Lenovo AIPCs, optimizing architecture and inference for sub-0.8-second generation; delivered ControlNet-based sketch/depth conditioning modules for B2B deployment.

SenseTime Research

May 2023 – Jan 2024

Research Intern, ISP & Codec Group

Beijing, China

- Proposed a task-aware encoder controller for visual-controllable video compression (CVPR 2024), using dynamic mode prediction and adaptive GoP selection to adapt pre-trained deep video codecs to detection, tracking, and action recognition, improving downstream bitrate-accuracy performance by over 25%.
- Proposed GaussianImage (ECCV 2024), the first image representation and compression framework based on 2D Gaussian splatting; optimized Gaussian parameterization/rendering and VQ-based quantization, achieving stronger low-bitrate RD performance with over 1,000 FPS decoding.

- Contributed to Boosting-NeRV, a universal conditional-decoder framework that improves reconstruction quality and convergence speed across multiple implicit video representation baselines (CVPR 2024).

SELECTED PUBLICATIONS

Salt: Self-Consistent Distribution Matching with Cache-Aware Training for Fast Video Generation

Xingtong Ge, Yi Zhang, Yushi Huang, Dailan He, Xiahong Wang, Bingqi Ma, Guanglu Song, Yu Liu, Jun Zhang.
ECCV 2026. [Paper] [Code] [Project Page]

AR-CoPO: Align Autoregressive Video Generation with Contrastive Policy Optimization

Dailan He, Guanlin Feng, **Xingtong Ge**, Yi Zhang, Bingqi Ma, Guanglu Song, Yu Liu, Hongsheng Li.
ECCV 2026. [Paper]

SenseFlow: Scaling Distribution Matching for Flow-based Text-to-Image Distillation

Xingtong Ge, Xin Zhang, Tongda Xu, Yi Zhang, Xinjie Zhang, Yan Wang, Jun Zhang.
ICLR 2026. [Paper] [Code] [Project Page]

Neighbor GRPO: Contrastive ODE Policy Optimization Aligns Flow Models

Dailan He, Guanlin Feng, **Xingtong Ge**, Yazhe Niu, Yi Zhang, Bingqi Ma, Guanglu Song, Yu Liu, Hongsheng Li.
CVPR 2026. [Paper]

LinVideo: A Post-Training Framework towards $\mathcal{O}(n)$ Attention in Efficient Video Generation

Yushi Huang, **Xingtong Ge**, Ruihao Gong, Chengtao Lv, Jun Zhang.
CVPR 2026. [Paper]

GaussianImage: 1000 FPS Image Representation and Compression by 2D Gaussian Splatting

Xinjie Zhang*, **Xingtong Ge***, Tongda Xu, Dailan He, Yan Wang, Hongwei Qin, Guo Lu, Jing Geng, Jun Zhang.
(*Equal contribution)
ECCV 2024. [Paper] [Code] [Project Page]

Task-Aware Encoder Control for Deep Video Compression

Xingtong Ge, Jixiang Luo, Xinjie Zhang, Tongda Xu, Guo Lu, Dailan He, Jing Geng, Yan Wang, Jun Zhang, Hongwei Qin.
CVPR 2024. [Paper]

Preprocessing Enhanced Image Compression for Machine Vision

Guo Lu, **Xingtong Ge**[‡], Tianxiong Zhong, Qiang Hu, Jing Geng. ([‡]Student first author)
IEEE TCSVT 2024 (arXiv 2022). [Paper] [Code]

HONORS & AWARDS

HKUST PGS Scholarship	2025
BIT Graduate Special Scholarship (twice)	2022, 2023
Outstanding Undergraduate Thesis Award (Top 5%)	2022
NAIC “Huawei Ascend Cup” AI+ Visual Feature Coding Competition (Top 2%)	2021–2022

TECHNICAL SKILLS

Programming	Python, C/C++, CUDA
ML & Systems	PyTorch, DDP, FSDP, profiling, Linux, Slurm/HPC, Git
Research	diffusion/flow models, distillation, post-training, controllable generation, neural compression