Shaozhe HAO

✉ shaozhe@connect.hku.hk   ⋄   ☎ (+852) 97100811 | (+86) 15997306891

🔗 https://haoosz.github.io/   ⋄   ⭘ https://github.com/haoosz

EDUCATION

The University of Hong Kong   Hong Kong, China
Ph.D. Candidate in Computer Science   *Sep. 2021 - early 2026*
 - *Supervisor: Prof. Kenneth K.Y. Wong and Prof. Kai Han*

Huazhong University of Science and Technology   Wuhan, China
B.Eng. in Automation,   GPA: 3.97/4.0 (Rank: 2/189)   *Sep. 2017 - Jun. 2021*

RESEARCH INTEREST

Generative Models, Diffusion Models, Visual Generation, Autoregressive Models, Representation Learning

WORK EXPERIENCES

Meta AI (Incoming)   Paris, France
Research Scientist Intern   *Aug. 2025 - Jan 2026*
 - Research on self-supervised learning with DINO team at Meta FAIR.

Tencent   Shenzhen, China
Machine Learning Intern   *Feb. 2025 - May 2025*
 - Supervised finetuned Qwen-2.5-VL-3B with LoRA by distilling Qwen-2.5-VL-32B for more finegrained video understanding of live short videos.
 - Developed retrieval-augmented generation (RAG) by incorporating transcriptions into prompts, using Whisper for automatic speech recognition (ASR) to process audio data.

Intellifusion Technologies   Shenzhen, China
Research Intern   Mentor: Dr. Xianbiao Qi   *Apr. 2024 - Dec. 2024*
 - Developed a novel conditional image generation model, BiGR, with 1.5B parameters, utilizing 32 A800 GPUs for DDP training. BiGR enables high-quality image generation following class conditions.
 - Extended BiGR to text-to-image generation. Curated a 24M dataset of image-caption pairs. Finetuned the model on the dataset with 32 A800 GPUs, which can perform 512×512 text-to-image generation.

SELECTED PUBLICATION

See my full publication list on Google Scholar.

[1] **Shaozhe Hao**, Xuantong Liu, Xianbiao Qi, Shihao Zhao, Bojia Zi, Rong Xiao, Kai Han, Kwan-Yee K. Wong. BiGR: Harnessing Binary Latent Codes for Image Generation and Improved Visual Representation Capabilities. ***ICLR***, 2025.
   - We introduce BiGR, a novel conditional image generation model using compact binary latent codes for generative training, focusing on enhancing both generation and representation capabilities.

[2] Xuantong Liu, **Shaozhe Hao**, Xianbiao Qi, Tianyang Hu, Jun Wang, Rong Xiao, Yuan Yao. Elucidating the Design Space of Language Models for Image Generation. ***ICML***, 2025.
   - We systematically investigate the design space of LLMs for image generation, analyze autoregressive models' learning behavior, and study the inherent differences between text and image modalities.

[3] Shuya Yang*, **Shaozhe Hao***, Yukang Cao, Kwan-Yee K. Wong. ArtiFade: Learning to Generate High-quality Subject from Blemished Images. *CVPR*, 2025. (*: equal contribution*)

  - We introduce ArtiFade, a novel approach designed to address the challenge of generating high-quality, subject-driven images from blemished image inputs.

[4] **Shaozhe Hao**, Kai Han, Zhengyao Lv, Shihao Zhao, Kwan-Yee K. Wong. ConceptExpress: Harnessing Diffusion Models for Single-image Unsupervised Concept Extraction. *ECCV*, 2024. (***Oral, top 2%***)

  - We present a novel diffusion-based method that can extract multiple instance-level concepts from a single image without any supervision like masks, concept words, or concept numbers.

[5] Shihao Zhao, **Shaozhe Hao**, Bojia Zi, Huaizhe Xu, Kwan-Yee K. Wong. Bridging Different Language Models and Generative Vision Models for Text-to-Image Generation. *ECCV*, 2024.

  - We present LaVi-Bridge, using LoRA and adapters to connect various pre-trained language models and generative vision models for text-to-image generation.

[6] **Shaozhe Hao**, Kai Han, Kwan-Yee K. Wong. CiPR: An Efficient Framework with Cross-instance Positive Relations for Generalized Category Discovery. *Transactions on Machine Learning Research* (**TMLR**), 2024.

  - We address generalized category discovery by exploiting cross-instance positive relations for contrastive learning, achieved by introducing a semi-supervised hierarchical clustering algorithm.

[7] **Shaozhe Hao**, Kai Han, Kwan-Yee K. Wong. Learning Attention as Disentangler for Compositional Zero-shot Learning. *CVPR*, 2023.

  - We address compositional zero-shot learning by using cross-attention to disentangle concept embeddings, while regularizing optimization with the earth mover's distance at the attention level.

HONORS & AWARDS

Arthur & Louise May Memorial Scholarship (HKU)	*2023*
Postgraduate Scholarships (HKU)	*2021-2025*
Outstanding Graduate (HUST)	*2021*
National Scholarship of China	*2018, 2019*

PROFESSIONAL SERVICE

Conference/journal reviewer for:

  - Conference on Computer Vision and Pattern Recognition (CVPR)
  - Conference on Neural Information Processing Systems (NeurIPS)
  - European Conference on Computer Vision (ECCV)
  - Association for the Advancement of Artificial Intelligence (AAAI)
  - Transactions on Machine Learning Research (TMLR)
  - Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
  - International Journal of Computer Vision (IJCV)
  - Transactions on Image Processing (TIP)

TA at HKU for:

  - COMP3317 Computer Vision (2022, 2023, 2024)
  - COMP7310 AI of Things (2023)

SKILLS

Programming Language	Python, Matlab, C/C++
Language	Mandarin (native), English (proficiency)
Others	PyTorch, Pandas, NumPy, Matplotlib, OpenCV, Git, LaTeX, Docker