Shaozhe Hao

szhao@cs.hku.hk
(852) 97100811 | (86) 15997306891
CB411, Chow Yei Ching Building, HKU
https://haoosz.github.io/

EDUCATION

The University of Hong Kong — Hong Kong SAR
Ph.D. Candidate in Computer Science — *Sep. 2021 - late 2025 or early 2026*
Supervisor: Prof. Kenneth K.Y. Wong and Prof. Kai Han

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

RESEARCH INTEREST

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

SELECTED PUBLICATIONS

See the full publication list on Google Scholar and the corresponding codes on GitHub.

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.** Submitted to *ICLR*, 2025. (*rating: 6,6,6,8,* ***top 15%***)

- 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.
- Project code: https://github.com/haoosz/BiGR

2 **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.
- Project code: https://github.com/haoosz/ConceptExpress

3 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.
- Project code: https://github.com/ShihaoZhaoZSH/LaVi-Bridge

4 **Shaozhe Hao**, Kai Han, Shihao Zhao, Kwan-Yee K. Wong. **ViCo: Detail-Preserving Visual Condition for Personalized Text-to-Image Generation.** Submitted to *IJCV*.

- We introduce a novel, lightweight, plug-and-play method to integrate visual conditioning into personalized text-to-image generation. Our method leads to improved performance.
- Project code: https://github.com/haoosz/ViCo

5 **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. We introduce a semi-supervised hierarchical clustering algorithm to produce such relations.
- Project code: https://github.com/haoosz/CiPR

6 **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.
- Project code: `https://github.com/haoosz/ade-czsl`

EXPERIENCE

Intellifusion Technologies Shenzhen, CN
Research intern Mentor: Dr. Xianbiao Qi *Apr. 2024 - Now*

- 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.
- The work is submitted to ICLR 2025 with the rating of 6,6,6,8.

The University of Hong Kong Hong Kong SAR
Research assistant Advisor: Prof. Kenneth K.Y. Wong *Jul. 2020 - Aug. 2020*

- Developed effective models for masked face recognition, achieving 98.23% verification accuracy on the masked LFW benchmark.

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

- TA at HKU for:
  - COMP3317 Computer Vision (2022, 2023, 2024)
  - COMP7310 AI of Things (2023)
- Conference/journal reviewer for:
  - Conference on Computer Vision and Pattern Recognition (CVPR)
  - 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)

SKILLS

Programming Languages	Python, Matlab, C/C++
Tools	PyTorch, OpenCV, LaTeX