WENYI MO

E-mail: mowenyi@ruc.edu.cn | Phone: (86)15815489499 | Homepage | Google Scholar

EDUCATION

Renmin University of China   Beijing, China
M.S. in Artificial Intelligence   *Sep. 2022 – Jun. 2025 (Expected)*
- Advisor: Prof. Bing Su.

South China University of Technology   Canton, China
B.E. in Computer Science   *Sep. 2018 – Jun. 2022*
- GPA: 3.91 / 4.0; Rank: 3 / 169
- China National Scholarship (Top *1%*).

RESEARCH INTERESTS

My research interests primarily lie in multimodal learning, with a focus on enhancing user alignment in generative models and improving image controllability in text-based conditions. Recently, I have concentrated on generative models (such as autoregressive models and diffusion models) and their applications, including aligning individual preferences and text-to-image generation.

PUBLICATIONS

Adaptive Preference Learning for Personalized Image Generation with Vision-Language Understanding
In Submission to CVPR   *2024*
- **Wenyi Mo**, Tianyu Zhang, Yalong Bai, Jieqiong Liu, Bing Su, Biye Li, Ji-Rong Wen
- [paper]
- *TL;DR:* A Vision-Language Model framework for personalized image generation that uses latent preference prototypes to model shared and unique user preferences, improving accuracy in preference prediction.

Dynamic Prompt Optimizing for Text-to-Image Generation
*Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), Seattle, USA*   *Jun. 16 - 20, 2024*
- **Wenyi Mo**, Tianyu Zhang, Yalong Bai, Bing Su, Ji-Rong Wen, Qing Yang
- [paper] [code]
- *TL;DR:* A reinforcement learning-based method for prompt optimization to improve text-to-image generation quality and user alignment.

Uniform Attention Maps: Boosting Image Fidelity in Reconstruction and Editing
*Proc. IEEE/CVF Winter Conference on Applications of Computer Vision (**WACV**), Tucson, USA* *Feb. 28 - Mar. 4, 2025*
- **Wenyi Mo**, Tianyu Zhang, Yalong Bai, Bing Su, Ji-Rong Wen
- [paper] [code]
- *TL;DR:* A tuning-free image editing technique that enhances fidelity in diffusion-based models using uniform attention maps.

MetaMask: Revisiting Dimensional Confounder for Self-Supervised Learning
*Proc. Advances in Neural Information Processing Systems (**NeurIPS**), New Orleans, USA, **Spotlight** Nov. - Dec. , 2022*
- Jiangmeng Li*, Wenwen Qiang*, Yanan Zhang, **Wenyi Mo**, Changwen Zheng, Bing Su, and Hui Xiong.
- [paper] [code]

Supporting Vision-Language Model Inference with Causality-pruning Knowledge Prompt
Arxiv Preprint.   *2024*
- Jiangmeng Li*, **Wenyi Mo***, Wenwen Qiang, Bing Su, and Changwen Zheng.
- [paper] [code]

RESEARCH EXPERIENCE

Research Intern   Mar. 2024 – Present
University of California, Santa Cruz   *Remote*
  • Supervisor: Prof. Cihang Xie
  • Research focus: Vision-Language Learning

Research Intern   Jan. 2024 – Mar. 2024
ByteDance, Applied Machine Learning Group   *Shanghai, China*
  • Supervisor: Dr. Yongfei Liu
  • Research focus: Controlled Image Generation

Research Intern   Sep. 2023 – Jan. 2024
Du Xiaoman Technology   *Beijing, China*
  • Supervisor: Dr. Yalong Bai
  • Research focus: Text-to-Image with Diffusion Model

PROJECTS

Prompt Optimizing for Text-to-Image Generation   Sep. 2023 – Jan. 2024

  • Proposed the Prompt Auto-Editing (PAE) method to dynamically optimize text prompts in text-to-image generation using reinforcement learning.
  • Introduced a two-stage training process: initial fine-tuning followed by online reinforcement learning to automatically adjust prompt modifiers, effect ranges, and weights.
  • Outperformed baseline methods on multiple datasets, with significant improvements in Aesthetic Score, CLIP Score, and PickScore. Achieved an Aesthetic Score of 6.12 (0.05 higher than human performance) and a PickScore of 73.9%, surpassing human-written prompts by 1.4%.

Image Reconstruction and Editing using diffusion model   Jan. 2024 - Sep. 2024

  • Developed a tuning-free image editing method that enhances image reconstruction fidelity in diffusion-based models using uniform attention maps.
  • Proposed an adaptive mask-guided editing technique to ensure consistency and precision during editing tasks.
  • Achieved notable improvements in reconstruction on the CelebA-HQ dataset, with an SSIM of 0.839 and a reduced LPIPS of 0.041. On the PIE benchmark, demonstrated a 12.4% improvement in background consistency (measured by MSE) and a 1.4% enhancement in editing accuracy for target areas (measured by CLIP Score).

SELECTIVE SCHOLARSHIPS AND AWARDS

  • **China National Scholarship:** Awarded to the top *1%* in the School of Computer Science. 2019
  • **China National Encouragement Scholarship:** Awarded to the top *3%* in the School of Computer Science. 2021
  • **Renmin University of China Scholarship**, 2024

TEACHING EXPERIENCES

  • Teaching Assistant: RUC, Comprehensive Artificial Intelligence Design, 2023 Fall
  • Teaching Assistant: RUC, Artificial Intelligence and Python Programming, 2023 Summer

PAPER REVIEWS

  • **Conference Reviewer:** NeurIPS 2024, ICLR 2025, WACV 2025, AISTATS 2025.

TECHNICAL SKILLS

  • Languages: Python, C/C++, LaTeX
  • Frameworks: Pytorch