YU LEI (雷宇)

 +86-135-6890-2648   ✉ leiyu2648@gmail.com   🌐 torpedo2003.github.io   ⬡ github.com/torpedo2003

RESEARCH INTERESTS

Current Focus: Personalized & Safe AI Generation • Preference Optimization for Generative Models • World Models & Interactive Generation

Previous Research (2022-2024): Medical Image Segmentation • Weakly Supervised Learning

EDUCATION

Ph.D. in Computer Science and Technology   **Sep 2025 – Present**

Institute of Computing Technology, Chinese Academy of Sciences (ICT, CAS)   Beijing, China

Advisor: Prof. Liang Li

Research Focus: Human-aligned generative AI, preference learning, world models

Bachelor of Engineering in Artificial Intelligence   **Sep 2021 – Jun 2025**

Sichuan University, College of Computer Science   Chengdu, China

Advisor: Prof. Lituan Wang

Honors: National Scholarship, Baogang Scholarship (Top 6 undergraduates), Outstanding Graduate

RESEARCH EXPERIENCE

Research Intern   **Feb 2025 – Aug 2025**

TeleAI, China Telecom   Beijing, China

Advisors: Dr. Jinbin Bai & Dr. Kaidong Yu

- **Research Focus:** Personalized safety alignment for text-to-image diffusion models
- **Key Contributions:** Developed PSAlign framework enabling user-specific safety controls via cross-attention adapters. Created Sage dataset (1,000 virtual users) with diverse safety preferences
- **Publications:** First-author paper under review at TMLR 2025; co-author on UPO (CVPR 2026 submission)

Research Assistant   **Sep 2022 – Aug 2024**

Machine Intelligence Laboratory, Sichuan University   Chengdu, China

Advisors: Prof. Lituan Wang & Prof. Lei Zhang

- **Research Focus:** Weakly supervised medical image segmentation using scribble annotations
- **Key Work:** Developed PCLMix with uncertainty-guided pixel-level contrastive learning
- **Publications:** First-author paper published at ICIC 2024, cited 6 times

PUBLICATIONS

Under Review

[1] **Yu Lei**, Jinbin Bai, Qingyu Shi, Aosong Feng, Kaidong Yu. "Personalized Safety Alignment for Text-to-Image Diffusion Models." *Under review at TMLR*, 2025. [arXiv] [Code] [Project]

[2] Jinbin Bai, Aosong Feng, **Yu Lei**, Zhuoran Zhao, Kaidong Yu. "Unpaired Preference Optimization: Aligning Visual Generative Models with Scalar Feedback." *Submitted to CVPR*, 2026. [Paper]

Published

[3] **Yu Lei**, Haolun Luo, Lituan Wang, Zhenwei Zhang, Lei Zhang. "PCLMix: Weakly Supervised Medical Image Segmentation via Pixel-Level Contrastive Learning and Dynamic Mix Augmentation." *International Conference on Intelligent Computing (ICIC)*, 2024. [arXiv] [Code]

Preprints & Survey Papers

[4] Jinbin Bai, **Yu Lei**, Hecong Wu, Yuchen Zhu, Shufan Li, Yi Xin, Xiangtai Li, Molei Tao, Aditya Grover, Ming-Hsuan Yang. "From Masks to Worlds: A Hitchhiker's Guide to World Models." *arXiv preprint*, 2025. [arXiv] [GitHub]

SELECTED OPEN SOURCE PROJECTS

PSAlign • ⬡ M-E-AGI-Lab/PSAlign • Python, 8 stars, 1 fork
Personalized safety alignment for text-to-image diffusion models with comprehensive documentation.

PCLMix • ⬡ torpedo2003/PCLMix • Python, 8 stars
Weakly supervised medical image segmentation framework with uncertainty-guided contrastive learning.

Awesome World Models • ⬡ M-E-AGI-Lab/Awesome-World-Models • Markdown, 58 stars
Curated collection of world models research, accompanying survey paper.

HONORS & AWARDS

Outstanding Graduate • Sichuan University	2024–2025
National Scholarship • Ministry of Education, China	2023–2024
Baogang Scholarship (Top 6 undergraduates) • Sichuan University	2023–2024
First Prize, Global Final • Huawei ICT Competition	2023–2024

LANGUAGES

Chinese: Native • **English:** Fluent (CET-4 578, CET-6 538)

Full publication list and references available at torpedo2003.github.io