

YUEQIN YIN

☎ (+1)7738919680

✉ yinyueqin0314@gmail.com

🎓 [Google Scholar](#)

Education

The University of Texas at Austin

Ph.D. in Artificial Intelligence

Austin, TX

Sept. 2023 - June. 2028(expected)

◇ Advisor: Prof. [Mingyuan Zhou](#)

CASIA (Institute of Automation, Chinese Academy of Sciences)

M.S. in Computer Vision and Deep Learning

Beijing, China

Sept. 2020 - June. 2023

Dalian University of Technology

B.S. in Software Engineering

Dalian, China

Sept. 2016 - June. 2020

◇ Grade: 91.8/100

◇ 2020 Outstanding Graduate, Liaoning Province

Under Review

1. **Yueqin Yin**, Yaxi Li, Xin Liu, Xun Wang, Kaiqiang Song, Simin Ma, Shujian Liu, Sathish Reddy Indurthi, Haoyun Deng, Pengcheng He, Mingyuan Zhou, Song Wang. ***ContextCheck: Sentence-Level Faithfulness Verification with Context-Aware Disambiguation.***
 - ◇ Proposes **ContextCheck**, a sentence-level faithfulness verification framework that explicitly conditions on preceding sentences for context-aware disambiguation (e.g., pronoun and reference resolution).
 - ◇ Shifts verification from isolated, context-free claims to claims evaluated within the natural, evolving flow of discourse, achieving state-of-the-art results on both standard and context-dependent faithfulness verification benchmarks.

Preprint

1. **Yueqin Yin***, Zhendong Wang*, Yi Gu, Hai Huang, Weizhu Chen, Mingyuan Zhou. ***Relative Preference Optimization: Enhancing LLM Alignment through Contrasting Responses across Identical and Diverse Prompts.*** [\[Paper\]](#) [\[Code\]](#) (* denotes equal contribution)
 - ◇ Proposes **Relative Preference Optimization (RPO)**, a contrastive preference learning framework that aligns large language models (LLMs) using both identical and semantically related prompts.
 - ◇ Introduces a **contrastive weighting mechanism** based on prompt similarity, allowing the model to learn from a richer and more diverse set of preference data, including unpaired examples, and to capture more nuanced human preferences.
 - ◇ Demonstrates superior alignment performance across dialogue and summarization tasks, achieving higher GPT-4 win rates than DPO, IPO, and KTO on the Anthropic-HH, OpenAI Summarization, and AlpacaEval 2.0 benchmarks.
2. **Yueqin Yin***, Zhendong Wang*, Yujia Xie, Weizhu Chen, Mingyuan Zhou. ***Self-Augmented Preference Optimization: Off-Policy Paradigms for Language Model Alignment.*** [\[Paper\]](#) [\[Code\]](#) (* denotes equal contribution)
 - ◇ Proposes **Self-Augmented Preference Optimization (SAPO)**, an off-policy framework that aligns LLMs without relying on paired preference data or external reward models, dynamically updating training samples through self-play.
 - ◇ Employs a replay buffer and an Exponential Moving Average (EMA) model to generate stable, high-quality negative samples for continual preference learning, leading to stronger and more consistent alignment.
 - ◇ Achieves superior results on benchmarks such as the Open LLM Leaderboard, IFEval, MT-Bench, and AlpacaEval 2.0, outperforming DPO, ORPO, and SPIN baselines in both reasoning and conversational ability.
3. Yi Gu, Zhendong Wang, **Yueqin Yin**, Yujia Xie, Mingyuan Zhou. ***Diffusion-RPO: Aligning Diffusion Models through Relative Preference Optimization.*** [\[Paper\]](#)
 - ◇ Proposes **Diffusion-RPO**, which extends Relative Preference Optimization (RPO) to text-to-image diffusion models by contrasting both identical and semantically related prompt-image pairs for improved preference alignment.

- ◇ Introduces a step-wise contrastive weighting mechanism and a new Style Alignment evaluation task, achieving superior performance over Diffusion-DPO and SFT on Stable Diffusion 1.5 and XL across human preference and style alignment benchmarks.

Publications

1. **Yueqin Yin***, Shentao Yang*, Yujia Xie, Ziyi Yang, Yuting Sun, Hany Awadalla, Weizhu Chen, Mingyuan Zhou. ***Segmenting Text and Learning Their Rewards for Improved RLHF in Language Model.*** Accepted by *Transactions on Machine Learning Research (TMLR)*, 2025. [\[Paper\]](#) [\[Code\]](#) (* denotes equal contribution)
 - ◇ Proposes a **segment-level reward model** for RLHF that assigns rewards to semantically coherent text segments, mitigating the limitations of sparse sequence-level and excessively granular token-level approaches.
 - ◇ Incorporates location-aware reward normalization and within-segment reward interpolation to improve PPO stability, yielding consistent performance gains on AlpacaEval 2.0, Arena-Hard, and MT-Bench benchmarks.
2. Zhangchen Xu, Yang Liu, **Yueqin Yin**, Mingyuan Zhou, Radha Poovendran. ***KODCODE: A Diverse, Challenging, and Verifiable Synthetic Dataset for Coding.*** Accepted by *ACL 2025 Findings*; 🏆 *Best Paper Award at DataWorld @ ICML 2025*. [\[Paper\]](#) [\[Code\]](#) [\[Website\]](#)
 - ◇ Introduces **KODCODE**, a large-scale synthetic coding dataset comprising 209K verified problems with executable solutions and unit tests, designed to advance code LLM fine-tuning and evaluation.
 - ◇ Employs a **three-stage generation pipeline** that synthesizes diverse, challenging, and verifiable programming tasks, enabling fine-tuned models such as Qwen2.5-Coder-32B-Instruct to achieve state-of-the-art results on HumanEval(+), MBPP(+), BigCodeBench, and LiveCodeBench.
3. Yanbo Xu*, **Yueqin Yin***, Liming Jiang, Qianyi Wu, Chengyao Zheng, Chen Change Loy, Bo Dai, and Wayne Wu. ***TransEditor: Transformer-Based Dual-Space GAN for Highly Controllable Facial Editing.*** In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7683–7692, 2022. [\[Paper\]](#) [\[Code\]](#) [\[Website\]](#) (* denotes equal contribution)
 - ◇ Proposes **TransEditor**, a Transformer-based dual-space GAN that introduces a **Cross-Space Interaction** mechanism between disentangled latent spaces (P-space and Z-space), enabling controllable and semantically consistent facial attribute editing.
 - ◇ Develops a **dual-space editing and inversion framework** that supports editing in one or both latent spaces, allowing fine-grained control of structural (pose) and textural (style) attributes simultaneously.
 - ◇ Achieves state-of-the-art controllability and identity preservation on CelebA-HQ and FFHQ, outperforming prior methods such as StyleGAN2, StyleMapGAN, and DAT in both qualitative and quantitative evaluations.
4. Shiyue Cao, **Yueqin Yin**, Lianghu Huang, Yu Liu, Xin Zhao, Deli Zhao, Kaiqi Huang. ***Efficient-VQGAN: Towards High-Resolution Image Generation with Efficient Vision Transformers.*** In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 7368–7377, 2023. [\[Paper\]](#)
 - ◇ Proposes **Efficient-VQGAN**, a two-stage vector-quantized framework for high-resolution image generation that replaces global attention with a **local attention-based quantization model** using Swin Transformer blocks, significantly improving efficiency and reconstruction fidelity.
 - ◇ Introduces a **multi-grained attention mechanism** combining global semantic and local contextual interactions, along with a unified autoencoding-training and autoregressive-inference paradigm, achieving state-of-the-art fidelity and speed on FFHQ and ImageNet benchmarks.
5. **Yueqin Yin**, Lianghua Huang, Yu Liu, Kaiqi Huang. ***DiffGAR: Model-Agnostic Restoration from Generative Artifacts Using Image-to-Image Diffusion Models.*** In *Proceedings of the International Conference on Computer Science and Artificial Intelligence (CSAI)*, 2022. [\[Paper\]](#)
 - ◇ Proposes **DiffGAR**, a model-agnostic diffusion framework for restoring clean images from diverse generative artifacts, treating artifact removal as an image-to-image translation task via conditional diffusion modeling.
 - ◇ Designs mechanisms to synthesize four representative types of artifacts from GANs, autoregressive, and diffusion models, and introduces a **class-conditioned classifier-free guidance** scheme enabling both blind and non-blind restoration with a tunable quality-fidelity trade-off.
 - ◇ Demonstrates superior restoration performance on synthetic and real artifact datasets, outperforming prior methods such as GFPGAN, Real-ESRGAN, and EdiBERT in both fidelity and human preference evaluations.

Internship & Project Experience

Generative Adversarial Network (GAN) Models.

Mar. 2021 – Dec. 2021 / Advisors: [Qianyi Wu](#), [Wayne Wu](#)

Shanghai AI Laboratory
Research Intern, Shanghai, China

- ◇ **Talking Face Generation.** Developed a neural system that generates talking faces driven by audio signals. Based on the First Order Motion Model (FOMM), replaced the image-driven branch with an *MFCC-based motion encoder*, allowing dense optical flow fields to warp source features for synchronized lip and facial movement synthesis.
- ◇ **GAN-based Image Generation and Editing.** Conducted a comprehensive literature review of state-of-the-art GAN-based generation and editing methods, focusing on disentangled representation and latent space manipulation. Introduced a **cross-space interaction mechanism** between the P -space and Z -space of StyleGAN, achieving more precise and controllable attribute editing. Published one paper at **CVPR 2022**.

Diffusion Models.

Mar. 2022 – Mar. 2023 / Advisors: [Lianghua Huang](#), [Yu Liu](#)

Alibaba DAMO Academy
Research Intern, Beijing, China

- ◇ **Continuous Diffusion Models.** Reproduced the Latent Diffusion Model (LDM) and conducted an extensive literature review of conditional generation and editing techniques, including blended latent diffusion, gradient-based editing, and classifier-free guidance. Introduced the **generative artifact restoration task** along with two new datasets, and developed **DiffGAR**, a conditional diffusion framework that effectively restores clean images from artifact-contaminated inputs. Published one paper at **CSAI 2022**.
- ◇ **Discrete Diffusion Models.** Proposed an improved prior modeling approach for VQ-Diffusion by designing a **two-layer VQGAN structure** that expands the single-layer Markov chain into a hierarchical two-layer process. Incorporated an auxiliary supervision loss on 32×32 discrete tokens to enhance prior learning on 16×16 tokens with minimal computational overhead.
- ◇ **Transformer-Based Two-Stage Vector-Quantized Model.** Developed a more efficient two-stage vector-quantized framework with enhancements to both the quantization and generative stages, achieving higher image quality and faster inference. Designed a new **hybrid image generation pipeline** combining autoencoding-based training with autoregressive inference, further improving synthesis fidelity. Published one paper at **ICCV 2023**.

Large Language Model Alignment.

Dec. 2023 – May. 2024 / Advisor: [Mingyuan Zhou](#)

The University of Texas at Austin
Research Assistant, Austin, TX

- ◇ **Relative Preference Optimization (RPO).** Extended Direct Preference Optimization (DPO) by introducing **relative comparisons across identical and semantically related prompts**, enabling models to learn more diverse preference signals. RPO achieves strong alignment performance across dialogue and summarization benchmarks.
- ◇ **Self-Augmented Preference Optimization (SAPO).** Proposed an **off-policy preference optimization framework** that dynamically augments training data through a replay buffer and an Exponential Moving Average (EMA) model. SAPO produces diverse and high-quality negative samples, leading to more stable and scalable preference learning. It achieves state-of-the-art results on AlpacaEval 2.0, IFEval, and MT-Bench, outperforming DPO and SPIN baselines.

Large Language Model Alignment and Reasoning.

June. 2024 – May. 2025 / Mentors: [Yujia Xie](#), [Ziyi Yang](#)

Microsoft Research
Research Intern, Redmond, WA

- ◇ **Segment-Level PPO.** Proposed a **segment-level reward model** for RLHF that bridges sparse sequence-level and overly fine-grained token-level rewards by assigning feedback to semantically coherent text segments. Incorporates location-aware normalization and within-segment interpolation, which significantly improve PPO training efficiency and alignment quality across AlpacaEval 2.0, Arena-Hard, and MT-Bench benchmarks. This method has been accepted by **TMLR 2025**.
- ◇ **KODCODE: A Diverse and Verifiable Synthetic Dataset for Coding.** Introduced **KODCODE**, a large-scale synthetic coding dataset with 209K verified problems and unit tests designed to improve code LLM training and evaluation. Accepted by **ACL 2025 Findings** and awarded **Best Paper Award at DataWorld @ ICML 2025**. Models fine-tuned on KODCODE achieve state-of-the-art performance on HumanEval(+), MBPP(+), BigCodeBench, and LiveCodeBench benchmarks.

Hallucination Detection and Verification.

June 2025 – Aug. 2025 / Mentor: [Song Wang](#)

Zoom Video Communications, GenAI Research Group
Research Intern, Bellevue, WA

- ◇ **ContextCheck: Sentence-Level Faithfulness Verification with Context-Aware Disambiguation.** Proposes **ContextCheck**, a hallucination detection framework that performs sentence-level verification while conditioning on preceding sentences for context-aware disambiguation (e.g., pronoun and reference resolution). It shifts verification from isolated, context-free claims to claims evaluated within the natural flow of discourse, achieving state-of-the-art results on both standard and context-dependent faithfulness benchmarks. *Under review.*

Academic Service

Reviewer for **CVPR 2024**, **ICCV 2025**, **ICLR 2025**, and **ICLR 2026**.