

葉適穎 Shih-Ying Yeh

Personal Profile

- **Contact:** kohaku@kblueleaf.net
- **Personal Homepage:** <https://kblueleaf.net>
- **GitHub Profile:** <https://github.com/KohakuBlueleaf>
- **HuggingFace Profile:** <https://huggingface.co/KBlueLeaf>
- **Google Scholar:** <https://scholar.google.com/citations?user=XmW5oGIAAAAJ>

Education & Career History

- **Undergrad student** Computer Science, National Tsinghua University (2021 – Present)
- **ML Engineer PT** Apple Inc. (2024/07 – 2024/11)
- **Consultant** Element, Inc. (2023/09 – 2024/07)
- **Web Developer** Fungs Design (2021/08 – 2022/02)

Expertise

- Generative Models: Text-to-Image Models, LLM
- Parameter-Efficient Fine-Tuning
- Deep Learning
- Computer Vision
- Web development

Publications

1. **Shih-Ying Yeh**, Yu-Guan Hsieh, Zhidong Gao, Bernard B W Yang, Giyeong Oh, & Yanmin Gong (2024). Navigating Text-To-Image Customization: From LyCORIS Fine-Tuning to Model Evaluation. In *The Twelfth International Conference on Learning Representations*.
 - ICLR 2024 Poster
 - Related open-source projects got over 3500 stars on GitHub in total.
2. Yu-Guan Hsieh, Cheng-Yu Hsieh, **Shih-Ying Yeh**, Louis Béthune, Hadi Pour Ansari, Pavan Kumar Anasosalu Vasu, Chun-Liang Li, Ranjay Krishna, Oncel Tuzel, & Marco Cuturi. (2024). Graph-Based Captioning: Enhancing Visual Descriptions by Interconnecting Region Captions.
 - Collaborated with Apple Inc.
 - Has been submitted to related conference.
3. **Shih-Ying Yeh**, Sang-Hyun Park, Giyeong Oh, Min Song, & Youngjae Yu. (2024). TIPO: Text to Image with Text Pre-sampling for Prompt Optimization.
 - Originated as my undergraduate project (advisor: Shang-Hong Lai prof.). Later collaborated with Yonsei University's laboratories and OnomaAI company.
 - Related open-source models have achieved over 200k downloads, related open source projects have received over 300 stars on GitHub in total.
 - Has been submitted to related conference.

Open-Source Projects / Contributions / Models

1. **LyCORIS - Lora beYond Conventional methods, Other Rank adaptation Implementations for Stable diffusion.**
 - GitHub Link: <https://github.com/KohakuBlueleaf/LyCORIS>
 - Open-source implementation of paper: “Navigating Text-To-Image Customization: From LyCORIS Fine-Tuning to Model Evaluation.”
 - Besides the algorithms introduced by LyCORIS, this project also implemented other frontier PEFT method, such as: GLoRA, DoRA, Diag-OFT, BOFT, (IA)^3
 - Has been widely used by Text-to-Image model and LLM community: have been included in Stable-Diffusion-Webui, sd-scripts, SimpleTuner, PEFT
2. **PixelOE - Detail-Oriented Pixelization based on Contrast-Aware Outline Expansion.**
 - GitHub Link: <https://github.com/KohakuBlueleaf/PixelOE>
 - Contrast-Aware Outline Expansion: Use contrast (difference between median and mean value of this patch) as information to blend eroded and dilated images.
 - Contrast-based down sample: Select most significant value within the patch to apply down sample
 - Near SoTA quality of pixelized images, without any Neural Network.
3. **KGen - A System for Prompt Generation to Improve Text-to-Image Performance.**
 - GitHub Link: <https://github.com/KohakuBlueleaf/KGen>
 - Open-source implementation of paper “TIPO: Text to Image with Text Pre-sampling for Prompt Optimization.” and its prototype “DanTagGen”.
 - [z-tipo-extension](#): open-source implementation to utilize TIPO in sd-webui or ComfyUI.
4. **TIFI: Temporal Inpainting for Frame Interpolation via Latent Diffusion Models and Jigsaw Sampling**
 - GitHub Link: <https://github.com/KohakuBlueleaf/TIFI>
 - Use conventional method (such as optical flow) to generate initial frame for interpolated frame
 - Use AnimateDiff-based LDM to train residual LDM on interpolated frames.
 - Use jigsaw sampling to enable the ability to produce infinite long videos.
5. **(Model) Kohaku-XL: Open sourced anime-style text-to-image generative model**
 - HuggingFace: [Kohaku XL - a KBlueLeaf Collection](#)
 - Finetuned SDXL-based text-to-image model.
 - Have been widely used by t2i community or company.
 - Million-scale dataset (8.4M for Kohaku-XL Zeta)
 - Trained on Consumer level hardware
6. **(Model) DTG and TIPO: pretrained automatic prompt engineering model for t2i task**
 - HuggingFace: [TIPO - a KBlueLeaf Collection](#), [DanTagGen - a KBlueLeaf Collection](#)
 - Pretrained models to be utilized in KGen library.
 - 100~500M LLM trained from scratch with 10~50B token seen.

7. (Contribution) Stable-diffusion-webui

- GitHub Link: <https://github.com/AUTOMATIC1111/stable-diffusion-webui>
- Top10 contributor and collaborator of this open-source project.
- Over 140k star

Research Interests

- Generative Models
 - Small model, PEFT, Customization.
 - Image: Text-to-Image, strong control (sketch to image or AI drawing software)
 - Text: General LLM, LM for transformation task, Agent System
 - Other: Multi-modal, Voice Clone TTS 、 AI Singer 、 Sheet Music Generation
- Representation Learning
 - Temporal, Spatial, Hierarchical
 - Real-world data
- Ranking Systems
 - Prediction of competition outcomes

Other Interests

- Literary, Visual Arts, Music
- Writing, Photography