

SHIRONG YANG

Generative AI | Computer Vision | Graphic Design

@ yangshr2024@shanghaitech.edu.cn Shanghai, China sloan-yang.com Sloan-Yang

RESEARCH FOCUS

My research focuses on **graphic design generation and structured visual generation**, including text-to-design, joint image-layout generation, content-edited layout generation, and SVG generation. I study how multimodal generative models coordinate visual appearance, spatial layout, typography, and editable structure, with broader interests in **2D/3D generation, world models, and generative modeling**.

EDUCATION

M.S. in Computer Science

ShanghaiTech University

2024 – Present Shanghai, China

GPA: 3.50/4.0

Research focus: Generative AI, Computer Vision, and Graphic Design Generation.

B.S. in Artificial Intelligence

Sichuan University

Sep 2020 – Jul 2024 Chengdu, China

GPA: 3.67/4.0

HONORS & AWARDS

-  **National Encouragement Scholarship**
2020–2021
-  **Tang Zhongying Moral Education Scholarship**
Class of 2020
-  **Outstanding Communist Youth League Member**
Sichuan University, 2020–2021
-  **Outstanding Student Leader**
Sichuan University, 2020–2021

RESEARCH EXPERIENCE

Graduate Researcher

ShanghaiTech University

2024 – Present Shanghai, China

My research focuses on generative models for graphic design, particularly multimodal generation, structured visual generation, and the interaction between visual content and layout.

Prior-Informed Design: Text-to-Design Generation with Multimodal Priors

- Proposed **PriDe**, a data-efficient text-to-design framework that leverages frozen text-to-image models and large language models as multimodal priors for generating structured and editable graphic designs.
- Developed **residual generative modeling** and prior-feature injection in a learned design latent space, enabling the generator to reuse pretrained knowledge of visual appearance, color, layout, and typography rather than relearning such knowledge from limited layered design data.
- Demonstrated high-quality text-to-design generation using only **10K layered training designs**, while maintaining competitive generation quality under further reduced-data settings.

RESEARCH INTERESTS

- Multimodal Generation
- Structured Visual Generation
- Graphic Design Generation
- World Models
- Generative Modeling

TECHNICAL SKILLS

Programming

- Python
- C/C++
- Rust

Deep Learning

- PyTorch
- Transformers
- Hugging Face Diffusers
- Diffusion Models
- Flow Matching
- VAE
- LLMs

Vision & Multimodal

- CLIP
- DINO
- Vision-Language Models
- Multimodal Learning
- 3D Generation

Learning Interaction between Image and Layout Priors for Joint Image-Layout Generation in Design Templates

- Proposed **InterIL**, a joint diffusion framework for simultaneously generating background images and foreground layouts, addressing the limitations of conventional sequential image-layout generation pipelines.
- Connected pretrained and frozen image and layout diffusion backbones with a learnable **bidirectional communication module**, allowing the model to explicitly learn image-layout interaction while preserving strong single-modality priors.
- Developed training-free inference guidance for incorporating user preferences into generated design templates without retraining the generative model.

Content-Edited Layout Generation for Graphic Design

- Formulated **content-edited layout generation**, extending conventional content-aware layout generation by jointly editing an input background image and generating a foreground layout according to a textual design intention.
- Proposed **Search-Compare-Edit (SCE)**, an automatic data construction framework for synthesizing paired image-editing data and textual editing instructions from large-scale image collections.
- Developed a two-flow multimodal **flow-matching** framework that jointly performs background-image editing and layout generation, explicitly coupling visual content with spatial composition.

PUBLICATIONS & MANUSCRIPTS

Prior-Informed Design: Text-to-Design Generation with Multimodal Priors

Shirong Yang, Ying Cao.
Under review at *ACM Transactions on Graphics (TOG)*, 2026.

Learning Interaction between Image and Layout Priors for Joint Image-Layout Generation in Design Templates

Shirong Yang, Bo Yang, Ying Cao.
Under review at *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 2026.

Content-Edited Layout Generation for Graphic Design

Shirong Yang, Ying Cao.
Under review at *IEEE Transactions on Multimedia (TMM)*, 2026.

Graphics & Game Development

- Unity
- Unreal Engine
- Bevy
- Blender
- Real-Time Rendering

Research Tools

- Linux
- Git
- LaTeX
- Hydra
- Weights & Biases

LANGUAGES

- Chinese
- English