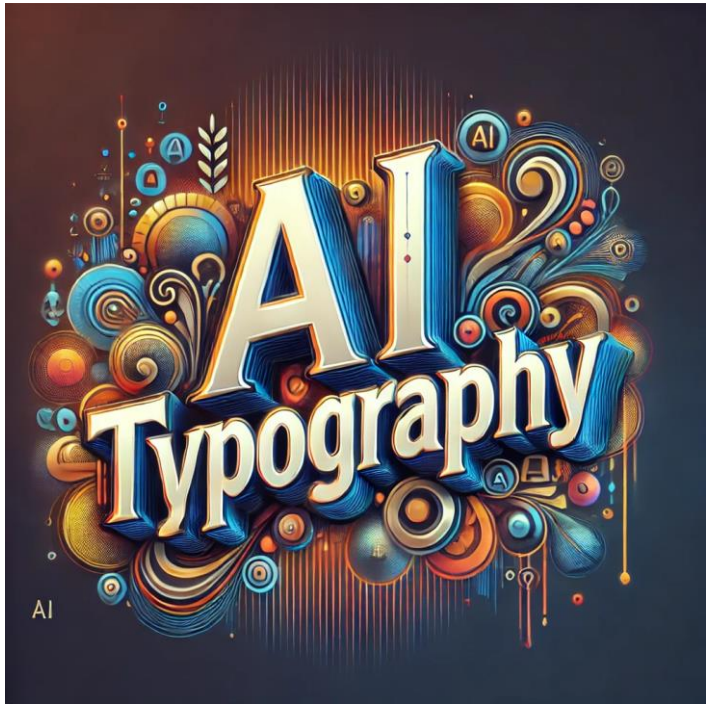
REVOLUTIONIZING TYPOGRAPHY: AI-DRIVEN METHODOLOGIES AND THEIR APPLICATIONS

30/7/2024

Presented by: Amin Bigdeli



UNIVERSITY OF WATERLOO | FACULTY OF MATHEMATICS

Introduction

- **Importance of AI in Typography**
  - **Enhanced Creativity**: AI enables designers to create complex and visually appealing text styles that were previously difficult to achieve.
  - **Efficiency**: Automates time-consuming tasks such as font synthesis and style transfer.
  - **Consistency and Quality**: Ensures high-quality output that maintains consistent style and aesthetics.

Introduction

- **Importance of AI in Typography**
  - **Enhanced Creativity**: AI enables designers to create complex and visually appealing text styles that were previously difficult to achieve.
  - **Efficiency**: Automates time-consuming tasks such as font synthesis and style transfer.
  - **Consistency and Quality**: Ensures high-quality output that maintains consistent style and aesthetics.
- **Objective of the Presentation**
  - **Methodology**: Techniques and algorithms developed for AI-driven Typography.
  - **Application**: Practical uses of these methodologies for different types of design.

Presentation Overview

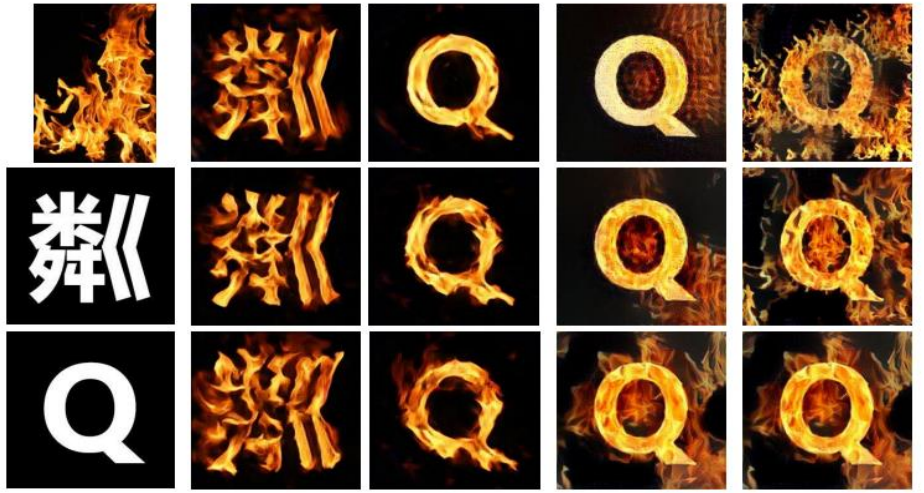
- Methodology
  - **Artistic and Dynamic Style Transfer:** Text Style Transfer, Font Style Transfer, and Font Style Interpolation
  - **Semantic Typography**
  - **User-Driven Typography**
- Applications
  - **Graphic and Visual Design:** Text Generation with AI, Logo Design, and Poster Design
  - **Scene Text Generation and Manipulation**

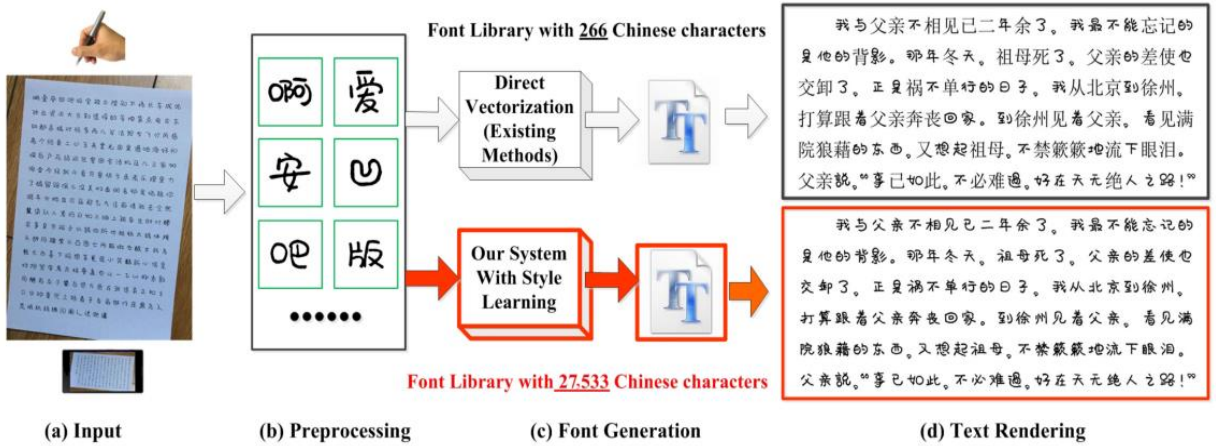
Methodology

Artistic and Dynamic Style Transfer

Changes the **Style** of a piece of **Text** based on the given Style.

- **Text Style Transfer**
  - Applying artistic effects and decorative elements to text.
  - Examples include adding textures, patterns, and other stylistic modifications to plain text to make it more visually engaging.
- **Font Style Transfer**
  - Transforming the overall style of font characters to create unique, artistic fonts.

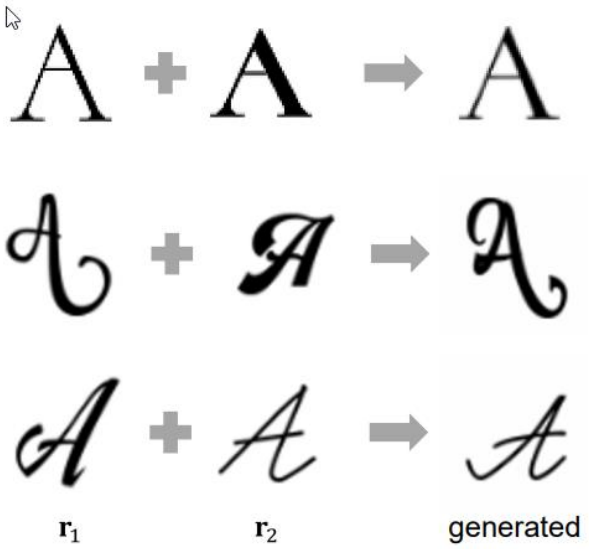




Yang et al. “Controllable Artistic Text Style Transfer via Shape-Matching GAN”, ICCV 2019
Lian et al. “EasyFont: A Style Learning-Based System to Easily Build Your Large-Scale Handwriting Fonts”, ACM Transactions on Graphic 2018

Artistic and Dynamic Style Transfer

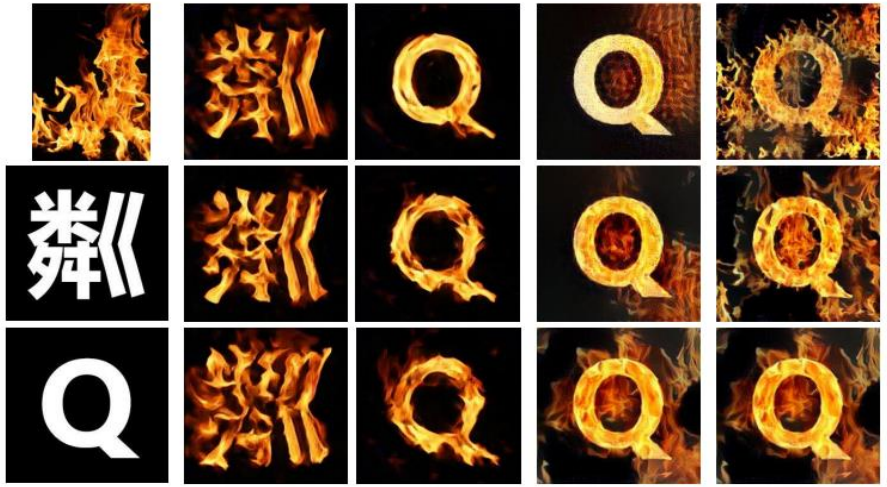
- **Font Style Interpolation**
  - Combines different font styles to create a smooth transition between them
  - Useful for creating hybrid fonts that merge characteristics of multiple styles, providing flexibility and variety in font design.



Kondo et al. “Font Style Interpolation with Diffusion Models”, arXiv 2024

Text Style Transfer - Paper Overview

- **Key Contributions**
  - **Novel Shape-Matching GAN Framework**
    - Bidirectional shape matching to map styles to glyphs at various deformation levels.
    - Real-time control of stylistic degree via an adjustable parameter.
  - **Scale-Controllable Module**
    - Continuous adjustment of glyph deformation.
    - Empowers the network to learn multi-scale shape features from style images and transfer them to target text.
  - **Real-Time Performance**
    - Achieves high-quality, diversified artistic text generation in real-time.



Yang et al. “Controllable Artistic Text Style Transfer via Shape-Matching GAN”, ICCV 2019

Text Style Transfer - Paper Overview

Methodology:

- **Backward Structure Transfer**: Simplifies the style image to various coarse levels, providing robust multi-scale shape mapping for data-driven learning.
- **Forward Structure Transfer**: Transfers shape features from the style image to the target text, achieving scale-controllable style transfer.
- **Texture Transfer Network**: Renders the texture in the style image onto the target text, completing the artistic text stylization process.

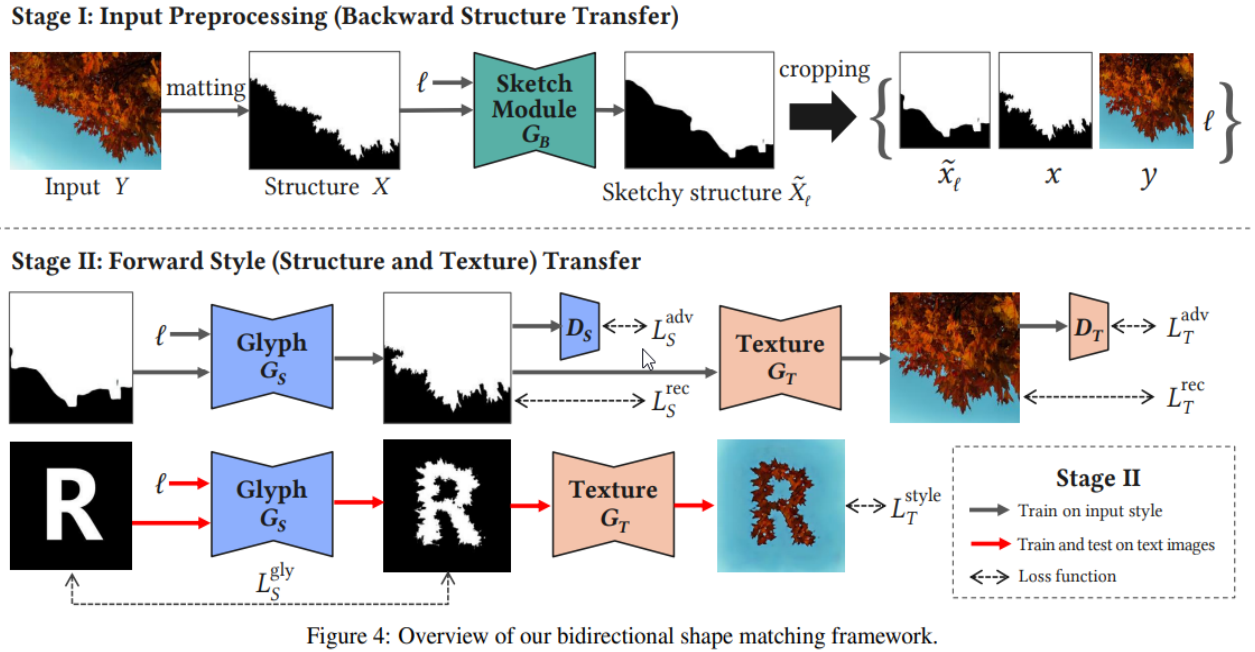


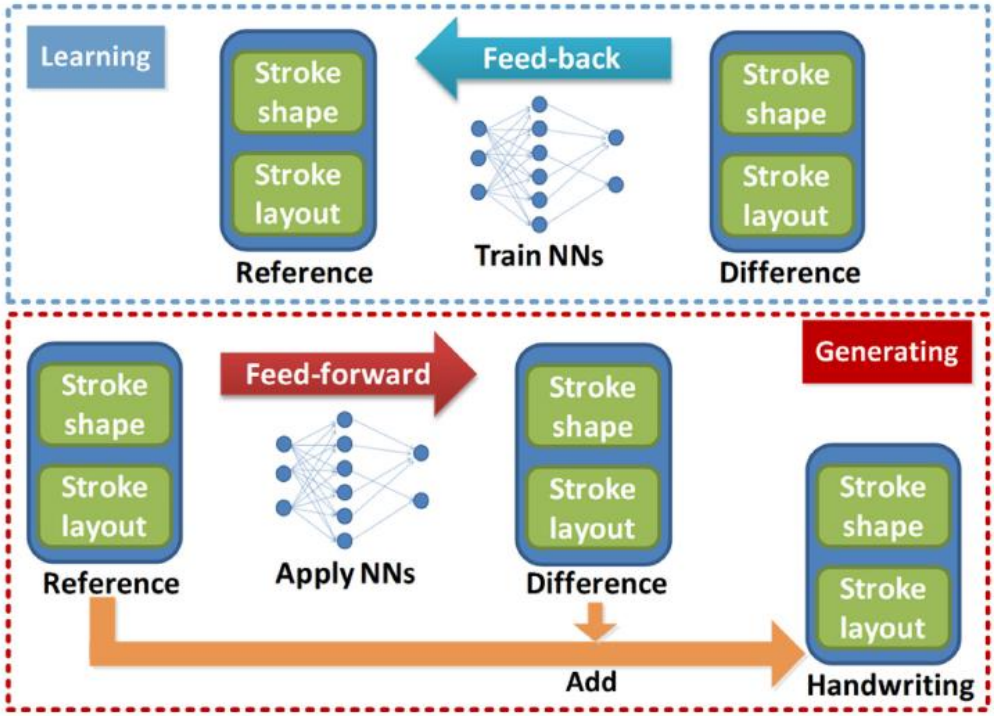
Figure 4: Overview of our bidirectional shape matching framework.

Yang et al. “Controllable Artistic Text Style Transfer via Shape-Matching GAN”, ICCV 2019

Font Style Transfer - Paper Overview

Methodology:

- **Stroke Extraction**
  - Utilizes a non-rigid point set registration approach to extract stroke trajectories from user-written samples.
  - Constructs a font skeleton manifold for reference.
- **Handwriting Style Learning**
  - Uses Artificial Neural Networks (ANNs) to learn stroke shape and layout styles.
  - Recovers handwriting details, such as stroke connectivity and contour shapes.
- **Handwriting Synthesis**
  - Generates new characters by applying learned style to reference data.
  - Combines human-written samples with machine-generated characters for better visual quality.



Lian et al. “EasyFont: A Style Learning-Based System to Easily Build Your Large-Scale Handwriting Fonts”, ACM Transactions on Graphic 2018

Font Style Transfer - Paper Overview

Dataset Statistics:

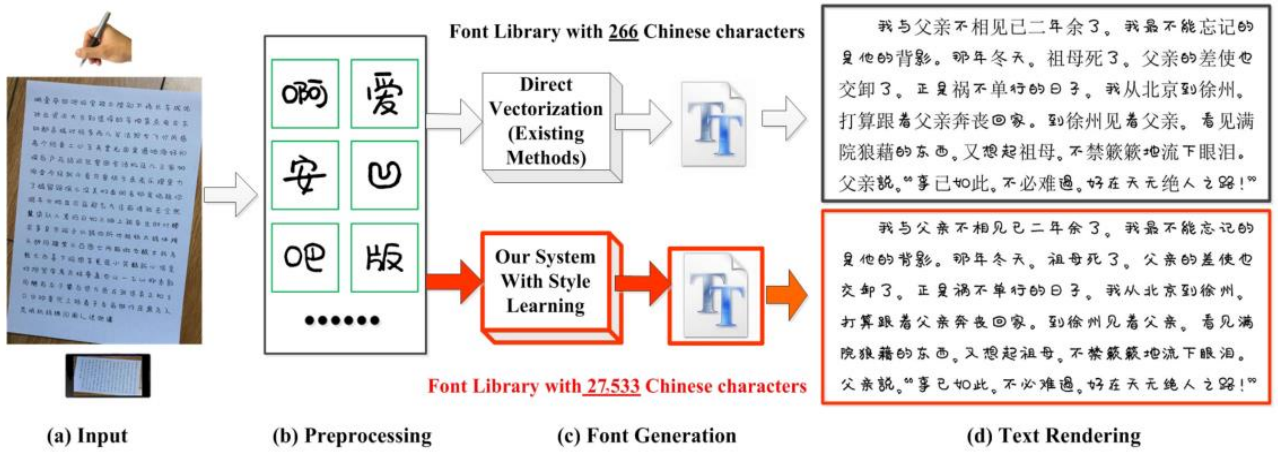
- **Training Data**:
  - Small set of carefully-selected samples (as few as 1% of the total characters) written by an ordinary person.
  - Includes multiple styles and variations to capture the full range of handwriting characteristics.
- **Character Set**:
  - Chinese characters: GB18030-2000 standard with 27,533 characters.
  - Adaptable to other writing systems with large character sets.

Lian et al. “EasyFont: A Style Learning-Based System to Easily Build Your Large-Scale Handwriting Fonts”, ACM Transactions on Graphic 2018

Font Style Transfer - Paper Overview

How it works:

- **Text Segmentation:** Obtain individual character images by segmenting rectified text pictures.
- **Stroke Extraction:** Extract the writing trajectory of each stroke for every character image.
- **Overall Style Learning:** Employ ANNs to learn the user's overall handwriting style.
- **Details Modeling:** Analyze and describe the connectivity of all sequential stroke pairs and details on the contour for each type of stroke.
- **Handwriting Synthesis:** Create trajectory for each character by adding the learned style on reference data and recovering handwriting details.
- **Font Generation:** Vectorize images of human-written samples and synthesis results for other characters.

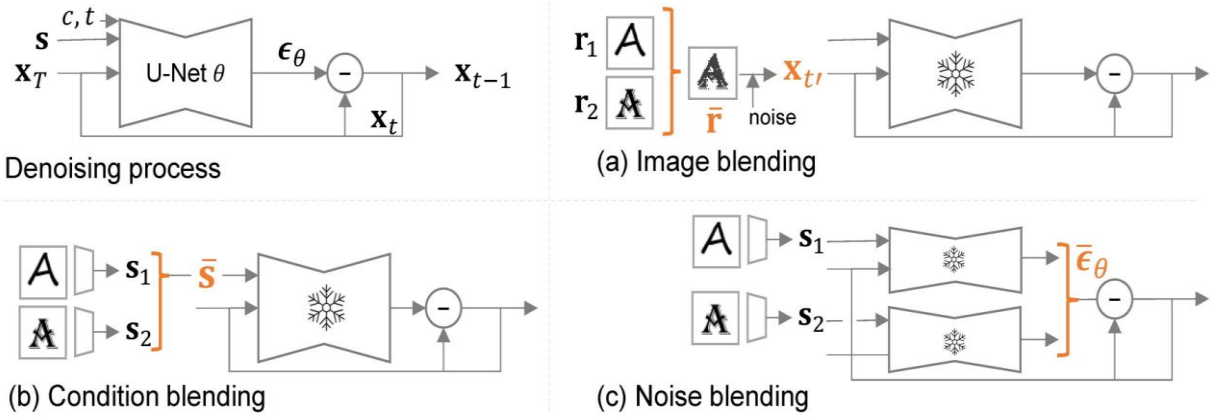


Lian et al. "EasyFont: A Style Learning-Based System to Easily Build Your Large-Scale Handwriting Fonts", ACM Transactions on Graphic 2018

Font Style Interpolation - Paper Overview

Methodology:

- **Conditional Diffusion Model**
  - Uses a U-Net architecture for denoising, conditioned on character class and style vector.
  - Generates character images by iteratively denoising from random noise.
- **Interpolation Approaches**
  - **Image-Blending**
    - Blends two character images using pixel-wise operations.
    - Generates a realistic image from the blended input during denoising.
  - **Condition-Blending**
    - Blends style feature vectors from two reference images.
    - Generates intermediate styles by interpolating style condition vectors.
  - **Noise-Blending**
    - Blends estimated noise images for two styles during the denoising process.
    - Affects the final generated image by merging styles at the noise level.



Kondo et al. "Font Style Interpolation with Diffusion Models", arXiv 2024

Font Style Interpolation - Paper Overview

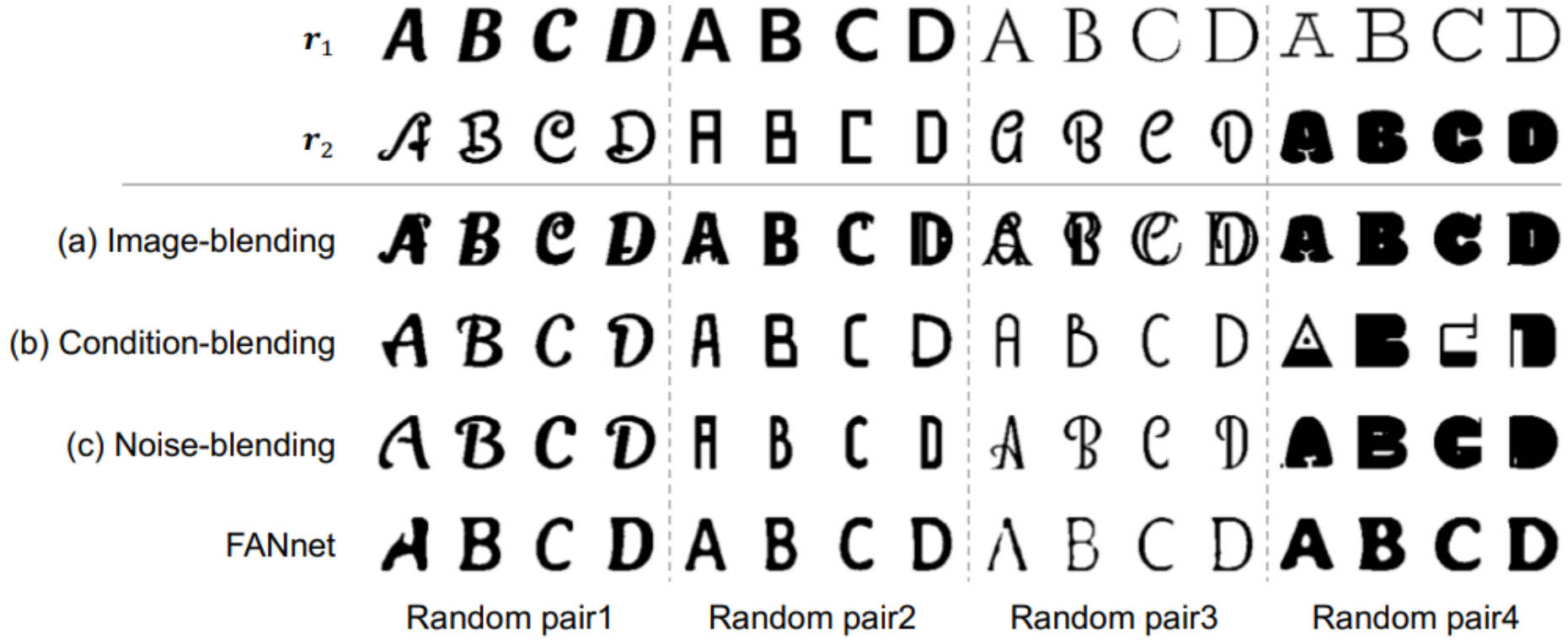
Dataset Statistics:

- **MyFonts Dataset**:
  - 17,412 fonts split into 13,938 training fonts, 1,734 validation fonts, and 1,740 test fonts.
  - Diverse styles including standard and decorative fonts.
- **GoogleFonts Dataset**:
  - Additional test set with less style diversity.
  - 2,545 fonts categorized into Serif, Sans-serif, Handwriting, and Display.
  - Includes font families with different weights (e.g., light, medium, bold).

Kondo et al. “Font Style Interpolation with Diffusion Models”, arXiv 2024

Font Style Interpolation - Paper Overview

Font Interpolation Examples



Kondo et al. “Font Style Interpolation with Diffusion Models”, arXiv 2024

Semantic Typography

What is it?

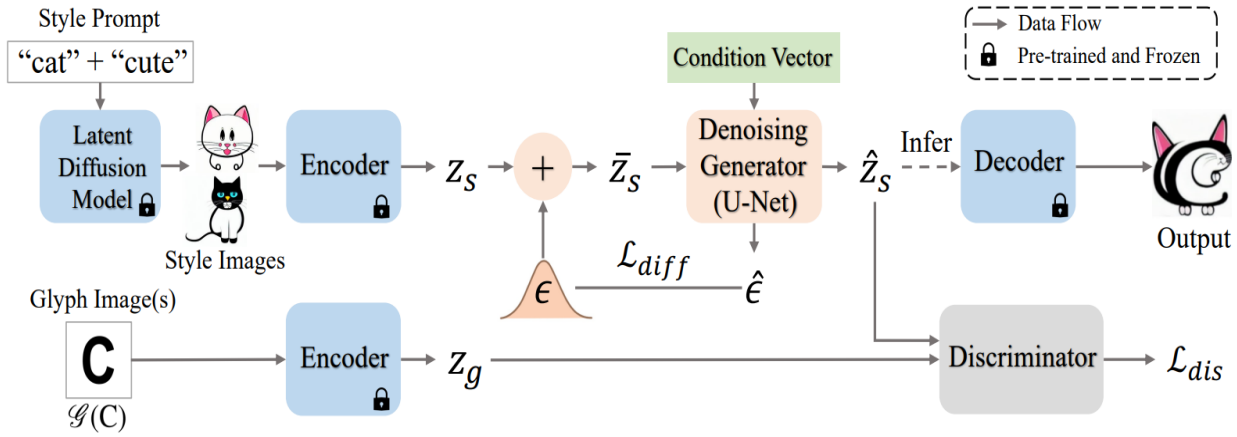
- **Definition**:
  - Semantic Typography integrates semantic understanding with typography design, allowing text to reflect its meaning and context visually.
- **Purpose**:
  - Enhance the expressiveness and functionality of text through intelligent and context-aware design.

Tanveer et al. “DS-Fusion: Artistic Typography via Discriminated and Stylized Diffusion”, ICCV 2023

Semantic Typography – Paper Overview

Methodology:

- **Latent Diffusion Model**
  - Utilizes a denoising generator to construct the latent space of given styles.
  - Encoder-decoder architecture conditioned on text prompts.
  - Gaussian noise applied to style images, denoised to produce the stylized glyph.
- **CNN-Based Discriminator**
  - Distinguishes between real and fake glyphs.
  - Ensures the stylized output retains the structure of the original font.
  - Loss Function: Combines diffusion loss and discriminator loss.



Tanveer et al. "DS-Fusion: Artistic Typography via Discriminated and Stylized Diffusion", ICCV 2023

Semantic Typography – Paper Overview

Dataset Statistics:

- **Font Datasets**
  - Variety of fonts collected from online repositories.
  - Includes serif, sans-serif, script, and decorative fonts.
  - Total of 50,000 font styles with diverse design characteristics.
- **Style Images**
  - Thousands of style images sourced from digital art databases.
  - Includes textures, patterns, and artistic elements.

Tanveer et al. "DS-Fusion: Artistic Typography via Discriminated and Stylized Diffusion", ICCV 2023

Semantic Typography

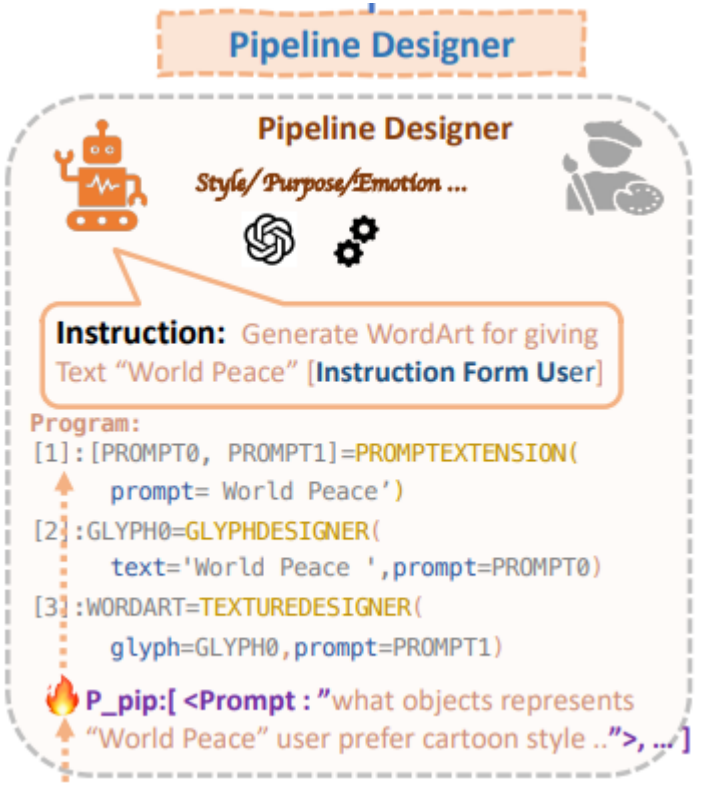
Impact on Typography

- **Enhanced Creativity**
  - **Artistic Expression**:
    - Enables designers to explore new creative horizons.
    - Facilitates the generation of unique and visually appealing typographic styles.
  - **Customization**: Allows for personalized font creation, catering to specific aesthetic preferences and branding needs.
- **Improved Readability and Semantics**
  - **Legibility**: Balances artistic stylization with readability, ensuring text remains clear and comprehensible.
  - **Semantic Relevance**: Incorporates semantic understanding to align visual style with the meaning and context of the text.
- **Broad Applicability**
  - **Advertising and Branding**: Provides tools for creating eye-catching and memorable typographic designs.
  - **Digital Art and Media**: Enhances the visual impact of digital art, posters, and multimedia content.

User-Driven Typography

Characteristics:

- **Active User Participation**:
  - Users are directly involved in decision-making processes.
  - They contribute ideas, preferences, and feedback throughout the design lifecycle.

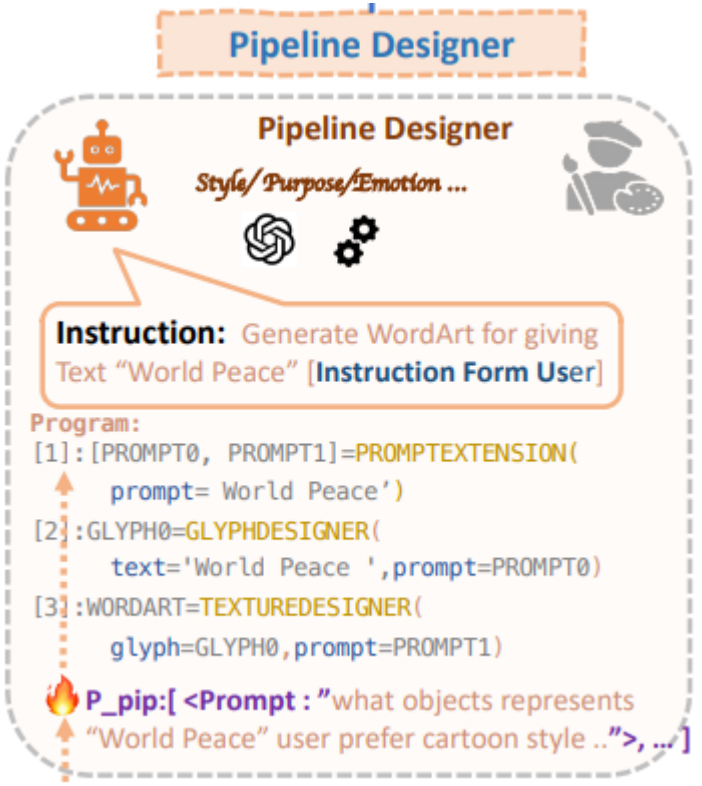


He et al. "MetaDesigner: Advancing Artistic Typography through AI-Driven, User-Centric, and Multilingual WordArt Synthesis", arXiv 2024

User-Driven Typography

Characteristics:

- **Active User Participation:**
  - Users are directly involved in decision-making processes.
  - They contribute ideas, preferences, and feedback throughout the design lifecycle.
- **Customizability:**
  - Design outcomes are highly customizable based on individual user inputs.
  - Users have significant control over the final product or service features.

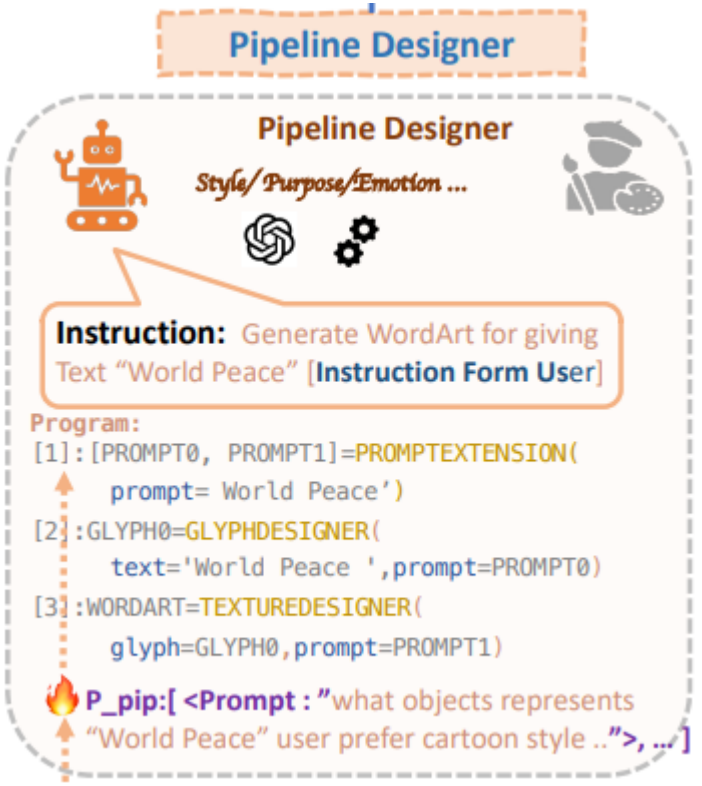


He et al. "MetaDesigner: Advancing Artistic Typography through AI-Driven, User-Centric, and Multilingual WordArt Synthesis", arXiv 2024

User-Driven Typography

Characteristics:

- **Active User Participation**:
  - Users are directly involved in decision-making processes.
  - They contribute ideas, preferences, and feedback throughout the design lifecycle.
- **Customizability**:
  - Design outcomes are highly customizable based on individual user inputs.
  - Users have significant control over the final product or service features.
- **Dynamic Interaction**:
  - Continuous interaction between designers and users.
  - Frequent iterations based on user feedback to refine the design.

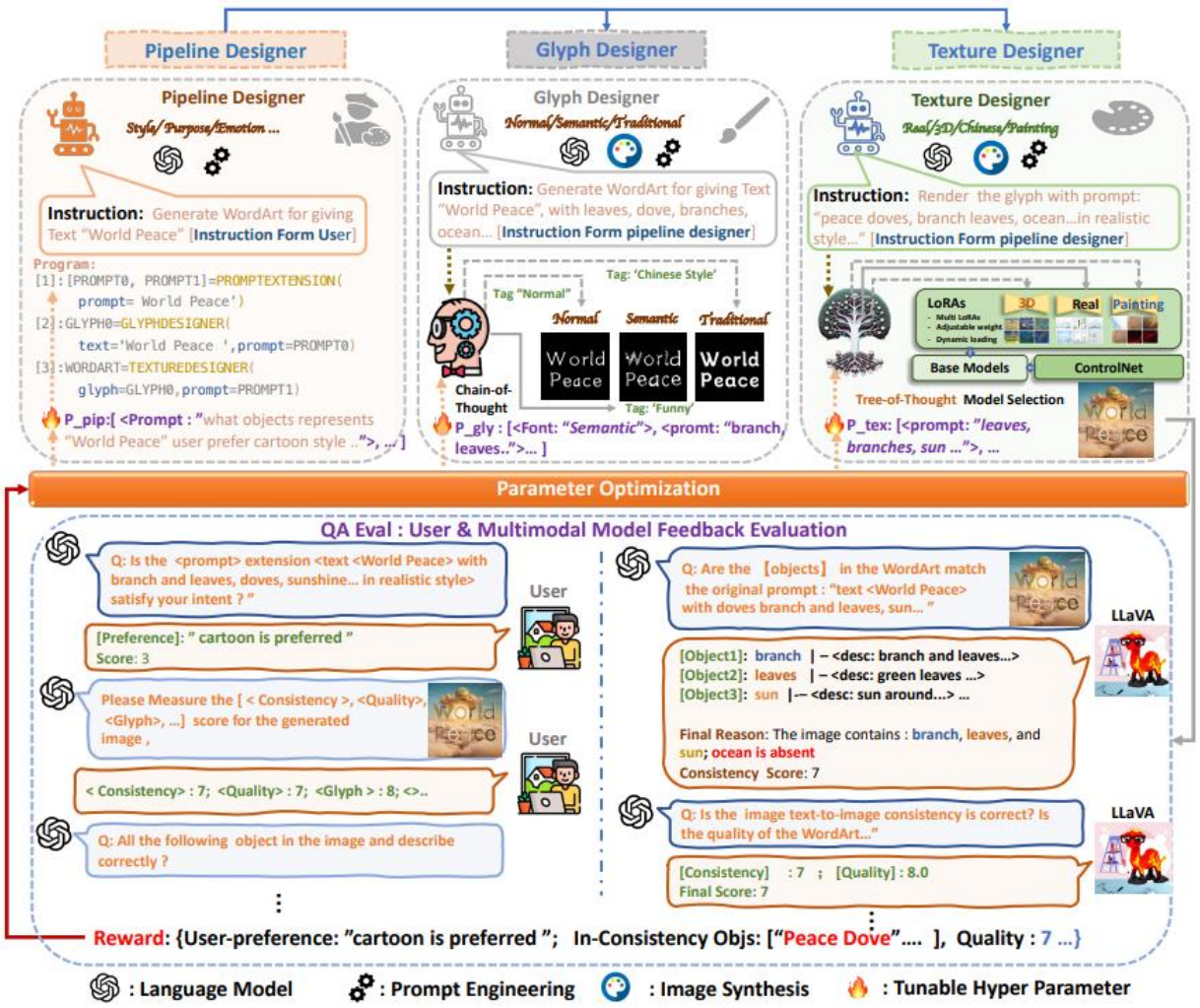


He et al. "MetaDesigner: Advancing Artistic Typography through AI-Driven, User-Centric, and Multilingual WordArt Synthesis", arXiv 2024

User-Driven Typography

Methodology Framework:

- **Pipeline Designer**:
  - Transforms user prompts into structured tasks for the Glyph and Texture Designers.
  - Uses GPT-4 for prompt extension and feedback integration.
- **Glyph Designer**:
  - Generates diverse glyph types (Normal, Traditional, Semantic) based on user inputs.
  - Ensures readability and context-appropriateness of the glyphs.
- **Texture Designer**:
  - Enhances glyphs with various texture styles using a Tree-of-Thought (ToT) model selection framework.
  - Integrates user feedback to refine texture designs.
- **Q&A Evaluation Agent**:
  - Iteratively refines the output based on multimodal and user feedback.
  - Adjusts hyperparameters to align with user-defined stylistic and thematic preferences.



He et al. "MetaDesigner: Advancing Artistic Typography through AI-Driven, User-Centric, and Multilingual WordArt Synthesis", arXiv 2024

UNIVERSITY OF WATERLOO | FACULTY OF MATHEMATICS

User-Driven Typography

Examples

"World Peace"

"世界和平"
"World Peace" in Chinese

"세계 평화" & "せかいへいわ"
"World Peace" in Korean and Japanese

He et al. "MetaDesigner: Advancing Artistic Typography through AI-Driven, User-Centric, and Multilingual WordArt Synthesis", arXiv 2024

Applications

Graphic and Visual Design

Objective: Enhance graphic and visual design by leveraging AI-driven typography techniques.

Significance: Enables the creation of visually appealing and personalized designs.

Key Applications:

- **Logo Design**: Creating Semantic Typographic Logos
  - Uses AI to generate logos that convey the semantic meaning of the brand or message.

Xiao et al. “TypeDance: Creating Semantic Typographic Logos from Image through Personalized Generation”, CHI 2024

Graphic and Visual Design

Objective: Enhance graphic and visual design by leveraging AI-driven typography techniques.

Significance: Enables the creation of visually appealing and personalized designs.

Key Applications:

- **Logo Design**: Creating Semantic Typographic Logos
  - Uses AI to generate logos that convey the semantic meaning of the brand or message.
- **Poster Design**: Multimodal Text Image Generation
  - Integrates visual harmony and text comprehension to create visually appealing posters.

"1人仅限1台"

Xiao et al. "TypeDance: Creating Semantic Typographic Logos from Image through Personalized Generation", CHI 2024
Gao et al. "TextPainter: Multimodal Text Image Generation with Visual-harmony and Text-comprehension for Poster Design", MM 2023

Graphic and Visual Design

- **Visual Text Generation**: Enhancing Text Rendering
  - Leverages language models to render text with enhanced visual appeal and context.
- **Applications**
  - **Branding and Advertising**: Custom logos and advertisements that stand out and effectively communicate the brand's message.
  - **Digital Media and Art**: Creating posters, digital art, and multimedia content with visually integrated and stylistically rich text.

Chen et al. "TextDiffuser-2: Unleashing the Power of Language Models for Text Rendering", arXiv 2023

Logo Design

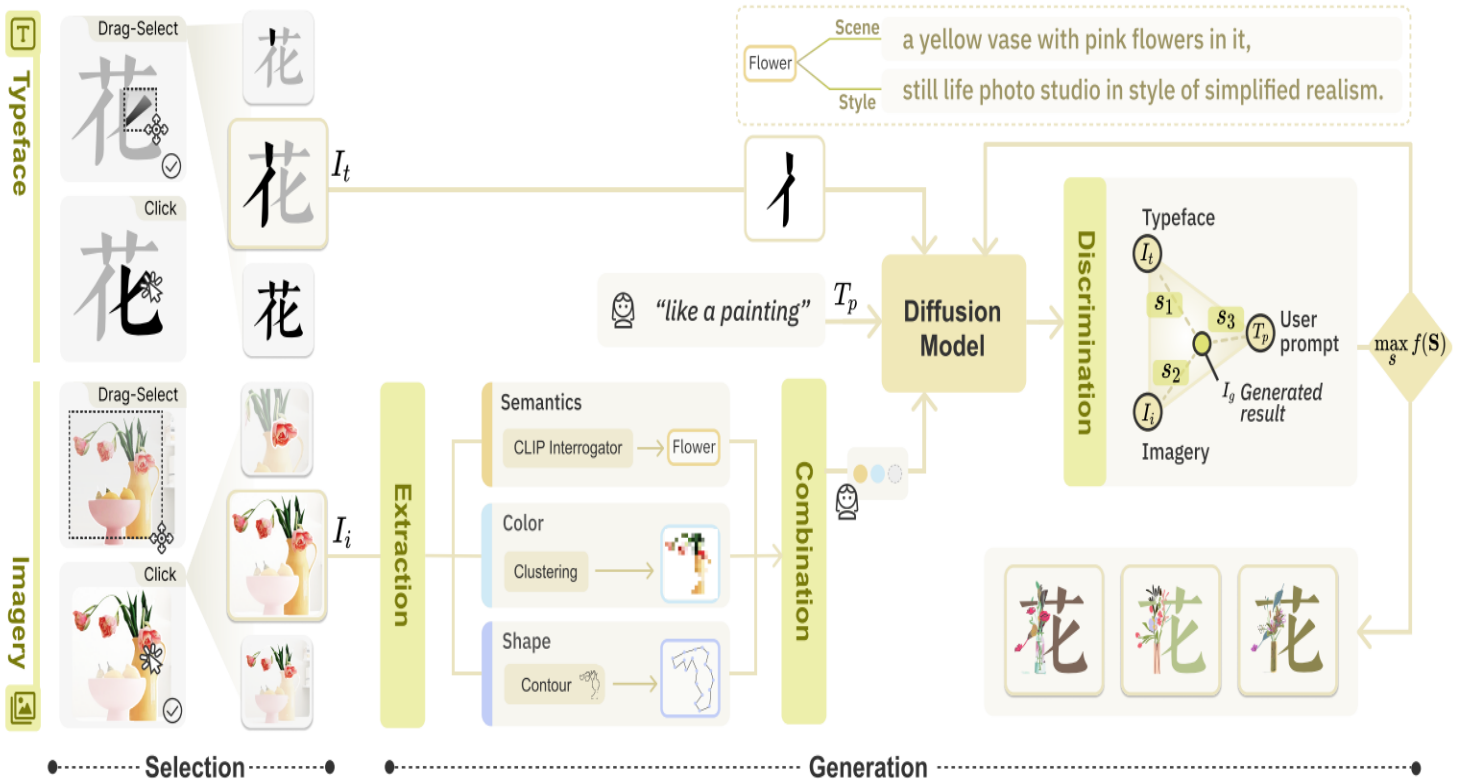
Model and Generation Process:

❖ Architecture

- Utilizes a diffusion model that combines image-to-image translation with text generation.
- Integrates semantic understanding to align typeface with imagery.

❖ Components

- **Encoder-Decoder Network**: Encodes the input image and decodes it into typographic elements.
- **StyleGAN**: Applies stylistic transformations to blend text and image elements.
- **Attention Mechanism**: Ensures the text integrates naturally with the image background.



Xiao et al. "TypeDance: Creating Semantic Typographic Logos from Image through Personalized Generation", CHI 2024

Poster Design

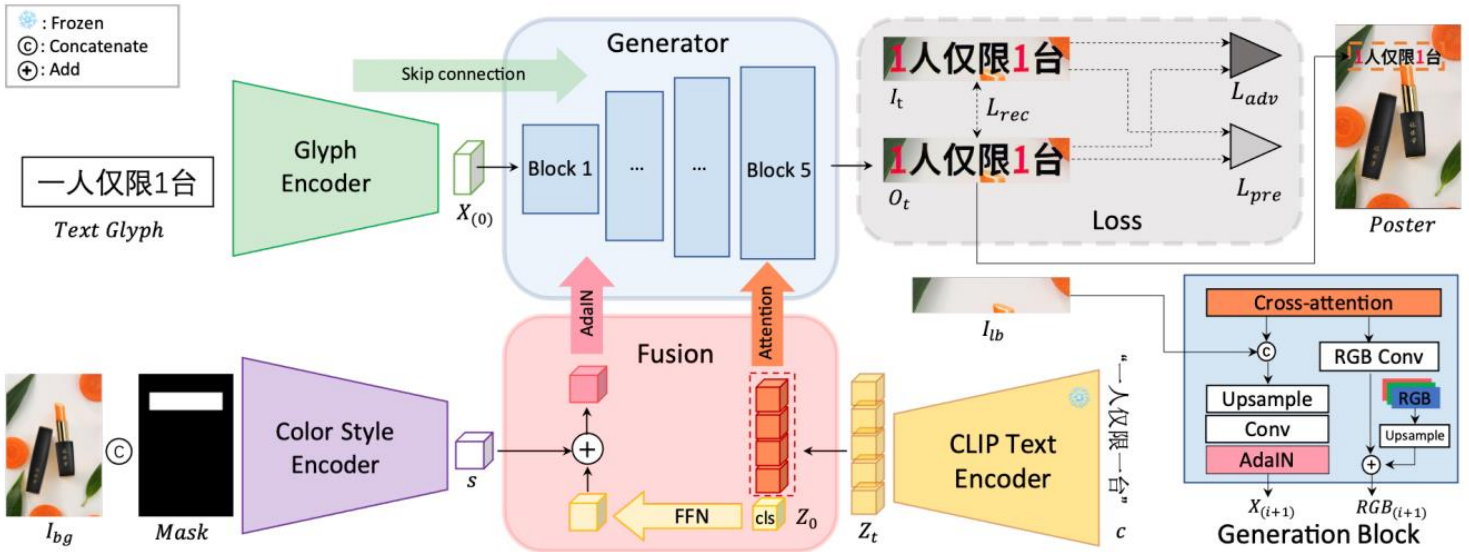
❖ Model

➢ Architecture

- Based on StyleGAN, incorporates glyph and style encoders.

➢ Components

- **Glyph Encoder**: Encodes text glyph features.
- **Color Style Encoder**: Extracts local and global color styles from the poster background.
- **CLIP Text Encoder**: Encodes text semantics at sentence and word levels.
- **Fusion Module**: Merges visual and semantic features.
- **Generator**: Generates text images that harmonize with the poster background.



Gao et al. "TextPainter: Multimodal Text Image Generation with Visual-harmony and Text-comprehension for Poster Design", MM 2023

Poster Design

❖ Generation Process

➢ Input

- User provides the poster background image, text content, and text position.

➢ Local and Global Color Harmony

- Color style encoder extracts style information from the poster background.

➢ Visual and Textual Fusion

- CLIP text encoder extracts semantic features from the text.
- Fusion module integrates these features with visual elements.

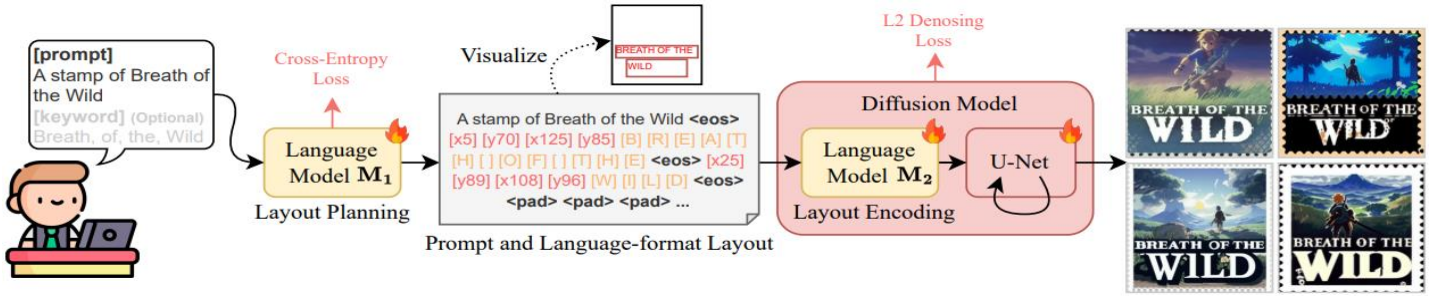
➢ Text Image Generation

- Generator produces the final text image, ensuring visual harmony and semantic relevance.

"1人仅限1台" ("only one for one person")
"1人仅限1台" ("only one for one person")
"1人仅限1台" ("only one for one person")
"1人仅限1台" ("only one for one person")

Gao et al. "TextPainter: Multimodal Text Image Generation with Visual-harmony and Text-comprehension for Poster Design", MM 2023

Visual Text Generation



- **Model**:
  - **Architecture**
    - Combines a fine-tuned large language model (LLM) with a diffusion model.
  - **Components**
    - **Layout Planning**: LLM plans text.
    - **Layout Encoding**: LLM encodes text position and content at the line level within the diffusion model.
- **Generation Process**
  - **Input:** User provides a text prompt and optionally keywords.
  - **Layout Planning**: LLM infers or uses provided keywords to determine text layout.
  - **Layout Encoding**: LLM encodes text layout into a format suitable for the diffusion model.
  - **Image Generation**: Diffusion model generates the final image with integrated text based on the encoded layout.

The handwritten words Hello World displayed on a wall in a neon light effect

A logo of Winter in artistic font, made by snowflake

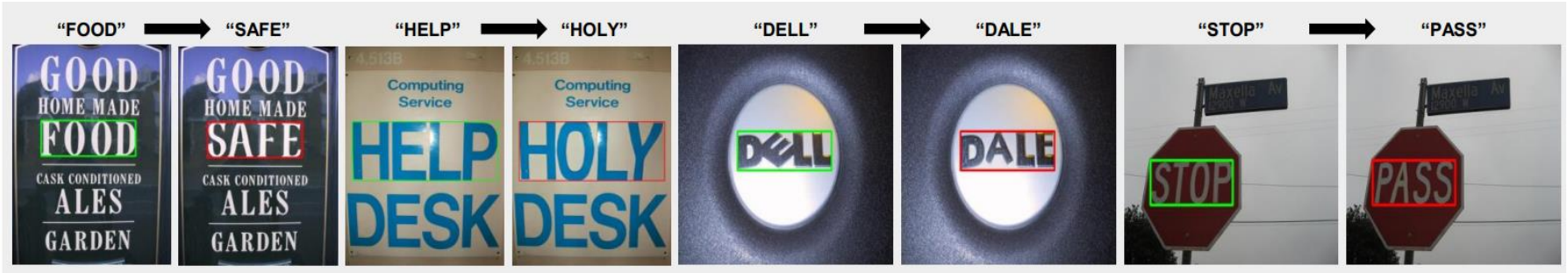
Chen et al. "TextDiffuser-2: Unleashing the Power of Language Models for Text Rendering", arXiv 2023

Scene Text Generation and Manipulation

- **Definition:** Creating and modifying text within images to appear naturally integrated into various backgrounds.
- **Application:** Digital media, advertising, augmented reality, and multilingual text rendering.

Generation

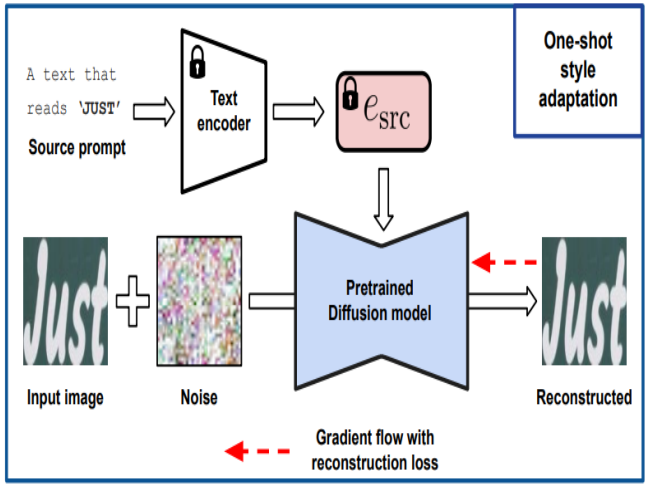
Manipulation

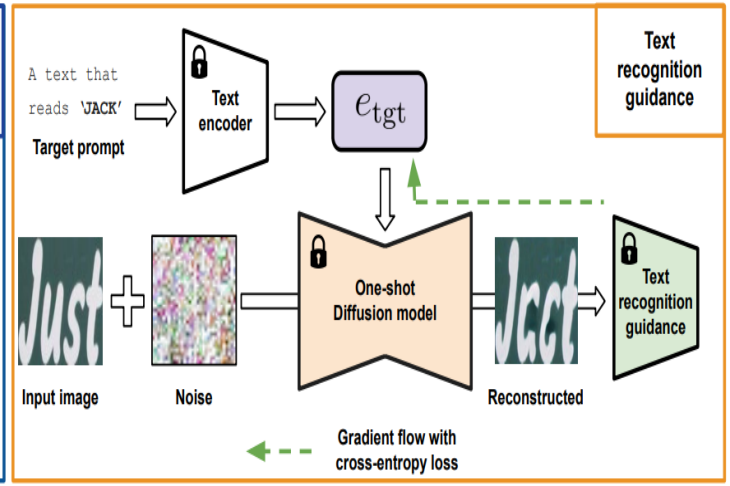


Santoso et al. "On Manipulating Scene Text in the Wild with Diffusion Models", WACV 2024
Zhang et al. "Brush Your Text: Synthesize Any Scene Text on Images via Diffusion Model", AAAI 2024

Scene Text Generation and Manipulation

Methodology:



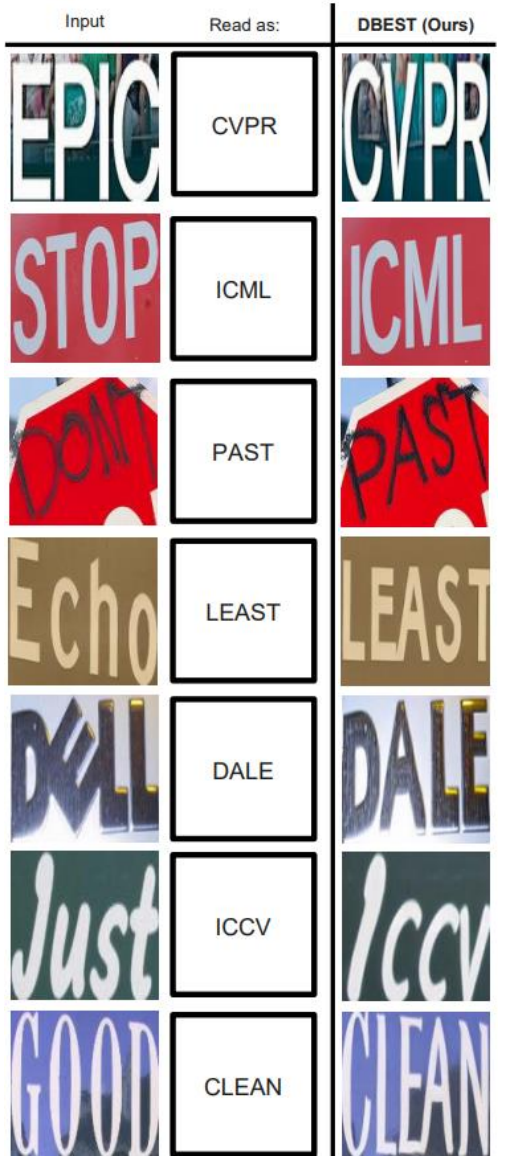


- **Input Processing**
  - Text is rendered into a sketch image and edge map.
  - Textual description of the scene is encoded as a prompt.
- **Diffusion Model**
  - Uses a Latent Diffusion Model (LDM) for denoising and image generation.
- **One-Shot Style Adaptation**
  - Fine-tunes the diffusion model to maintain the source style on the edited text.
- **Text Recognition Guidance**
  - Uses a text recognition model to guide the diffusion process, ensuring accuracy and readability.

Santoso et al. “On Manipulating Scene Text in the Wild with Diffusion Models”, WACV 2024

Scene Text Generation and Manipulation

Examples



Santoso et al. "On Manipulating Scene Text in the Wild with Diffusion Models", WACV 2024

Conclusion

Key Insights

❖ Advancements in AI-Driven Typography

- AI models, such as diffusion models and GANs, significantly improve the creativity and functionality of text and font generation.
- Integration of user inputs and preferences enhances personalization and engagement in design processes.

❖ Impact on Visual Communication

- AI-driven techniques facilitate the creation of visually appealing and contextually relevant designs.
- Applications span across branding, digital media, advertising, and educational content, offering versatile tools for designers.

❖ Future Directions

- Continued development in AI and machine learning will further refine and expand the capabilities of typography and visual design tools.
- Potential for more interactive and user-friendly design platforms, leveraging real-time feedback and advanced semantic understanding.

Thank you!
Any Questions?

Amin Bigdeli abigdeli@uwaterloo.ca