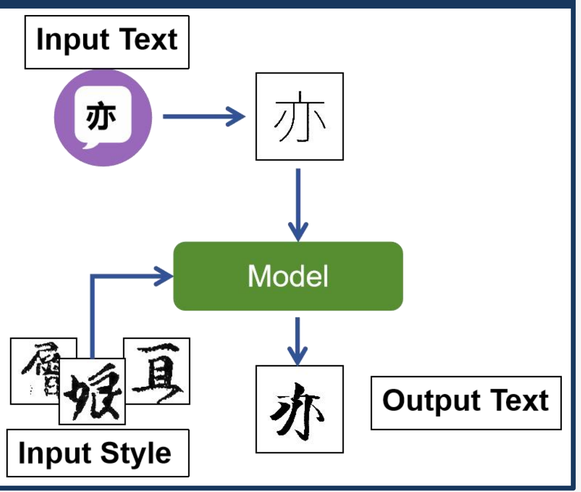
OS1D-10

Few-shot Font Generation for Japanese Kuzushiji with Differentiable Renderer

Honghui Yuan, Keiji Yanai
The University of Electro-Communications

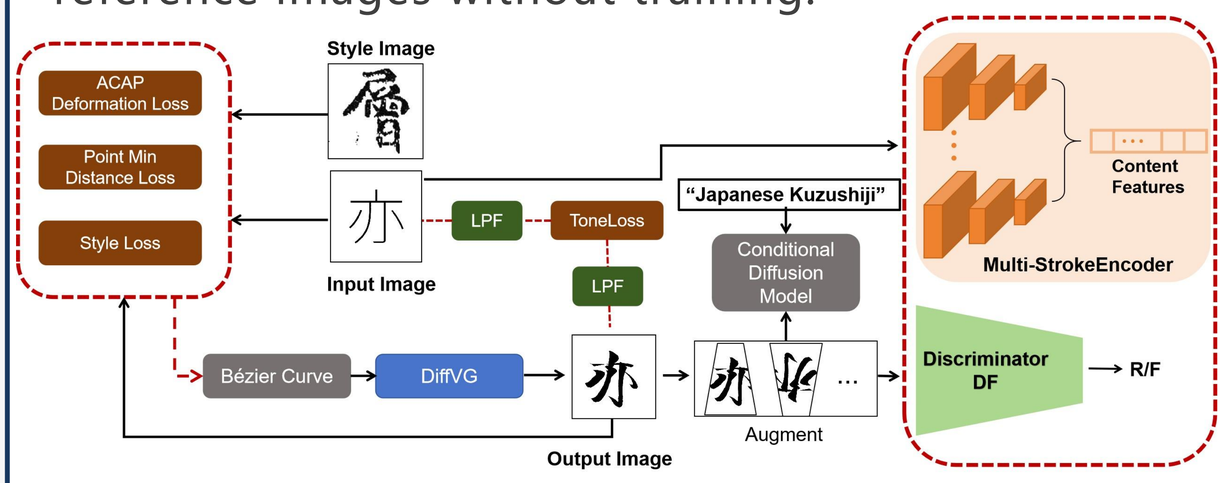
Abstract

- Transfer modern font to **Kuzushiji** font using Diffusion and DiffVG.
- Generate hand-written characters using **vector-based** images.



Methodology

- We propose a novel model based on WAI[1] to **generate vector Kuzushiji characters** using a few reference images without training.



Network

- **Conditional Stable Diffusion model**
- **DiffVG**
- **Discriminator**
- **Multiple-stroke encoder**

Main Loss Functions

1.Style Loss: mean squared error from style images

$$L_{\text{style}} = \sum \|G_{l,i,j}(x_s) - G_{l,i,j}(x_g)\|^2$$

$$G_l(x) = F_l(x)F_l^T(x)$$

2.Multi-stroke ACAP Loss:
Calculate the angle for each stroke

$$\mathcal{L}_{acap}(\mathbf{P}, \hat{\mathbf{P}}) = \frac{1}{k}\sum_{j=1}^{k}\left(\sum_{i=1}^{m_j}(\alpha_j^i - \hat{\alpha}_j^i)^2\right)$$

3.SDS loss [1]:
Convey semantic concepts from the prompt to the images

$$\nabla_\theta \mathcal{L}_{LSDS} = \mathbb{E}_{t,\epsilon}\left[w(t)\left(\hat{\epsilon}_\phi(\alpha_t x_t + \sigma_t\epsilon, y) - \epsilon\right)\frac{\partial z}{\partial z_{aug}}\frac{\partial x_{aug}}{\partial \theta}\right]$$

4.Point Min Distance Loss:
Make the distances of each point change gradually

Results

1.Kuzushiji character generation from source modern font characters

2. Comparison

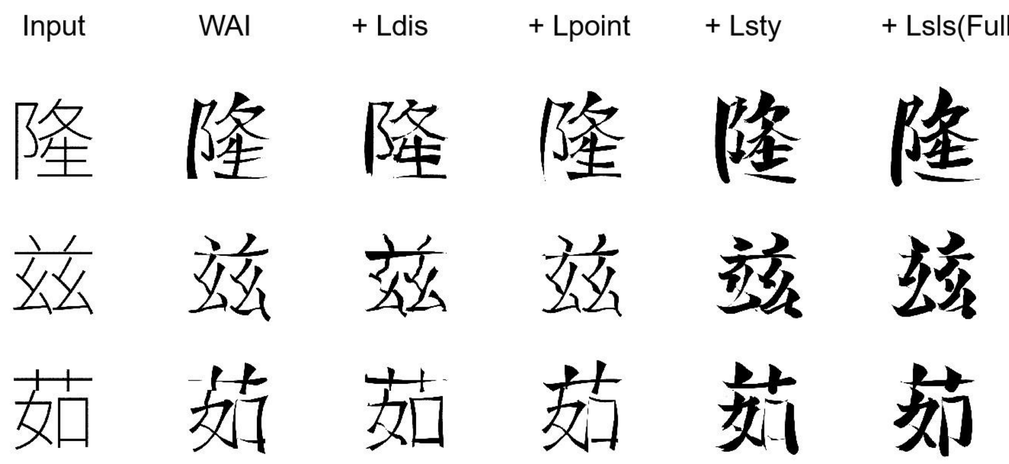
Generates cleaner and more structurally accurate glyphs

Our results have achieved an improvement of more than 6% in LPIPS, SSIM, and FID

	LPIPS↓	SSIM↑	FID↓	RMSE↓
CLIPFont	0.4137	0.5002	1.8965	**0.4318**
WordAsImage	0.3903	0.4887	1.8531	0.4498
Ours	**0.3634**	**0.5592**	**1.5494**	0.4379
	6.89%	11.79%	16.38%	(-)

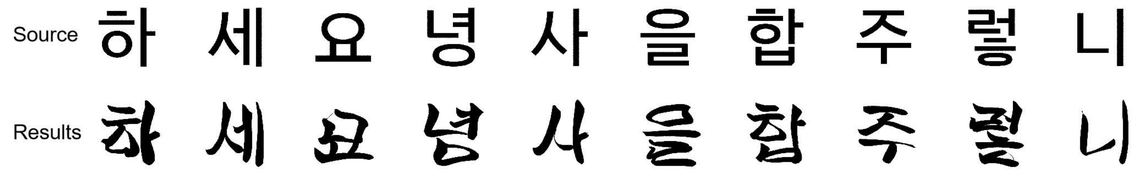
3.Ablation Studies

Full model can generates cleaner and more structurally accurate glyphs

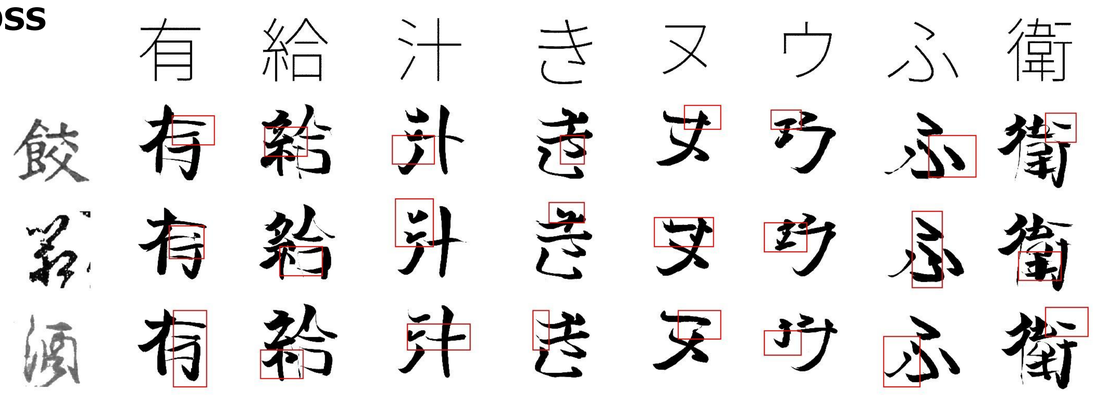


4. Other Results

The performance of our model in other languages



Using different reference style images used in the style loss



[1] Iluz, Shir, et al. "Word-as-image for semantic typography." ACM Transactions on Graphics (TOG) 42.4 (2023): 1-11.