KuzushijiDiffuser: Japanese Kuzushiji Font Generation with FontDiffuser

Honghui Yuan, Keiji Yanai

The University of Electro-Communications, Tokyo, Japan

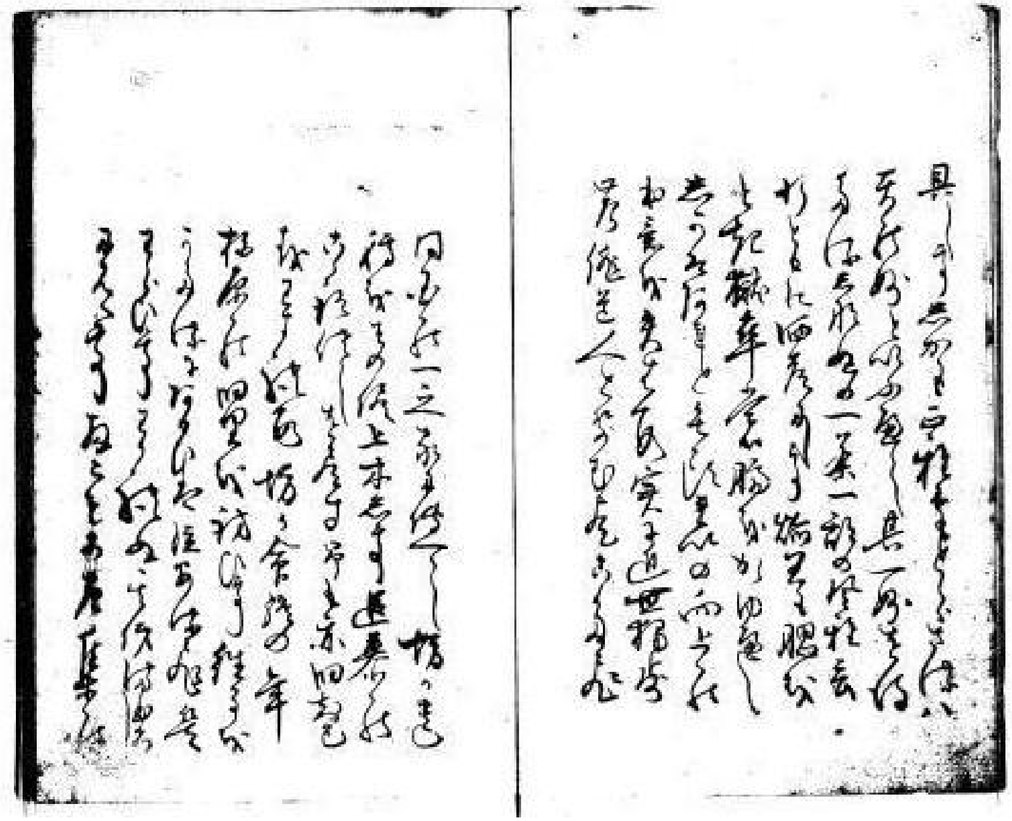
Introduction

- Image generation using deep learning has made remarkable progress.
- Various methods for generating fonts have been developed with great success.
- Reduce the time-consuming and labor-intensive process in designing fonts.

Chen, Xinyuan, et al. "DGFont++: robust deformable generative networks for unsupervised font generation." arXiv preprint arXiv:2212.14742 (2022).

Introduction

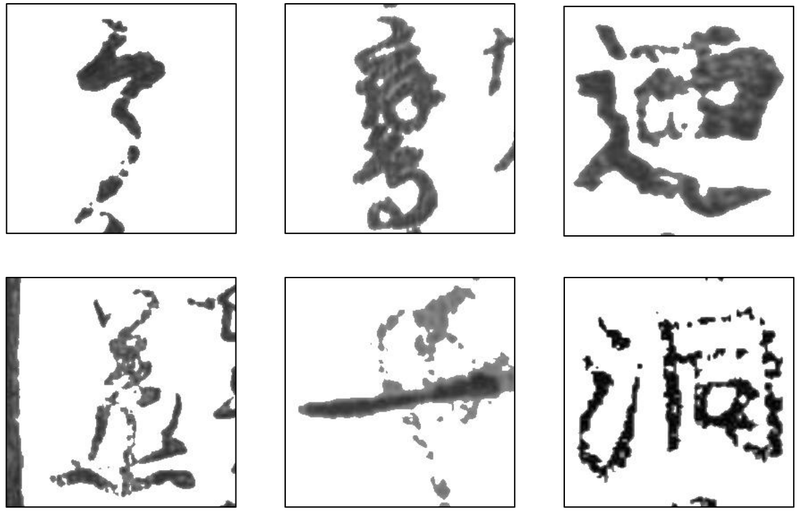
- Kuzushiji characters were used in Japan hundreds of years ago.
- Many valuable ancient documents are written in Kuzushiji.



Cited from http://codh.rois.ac.jp/

Problem

- The deteriorated historical documents leading to some characters being lost or illegible.
- Kuzushiji have more complex structures than modern fonts.



Kuzushiji Character Samples

Modern Font　　Kuzushiji

Problem

- The identification and generation of Kuzushiji are more challenging than that of modern font characters.
  1. Different characters may have similar forms.
  2. Same character have different forms.

Require more detailed analysis of their characteristics

Problem

- Existing methods are limited to modern character generation.
- Few studies have been done on handwritten complex fonts.

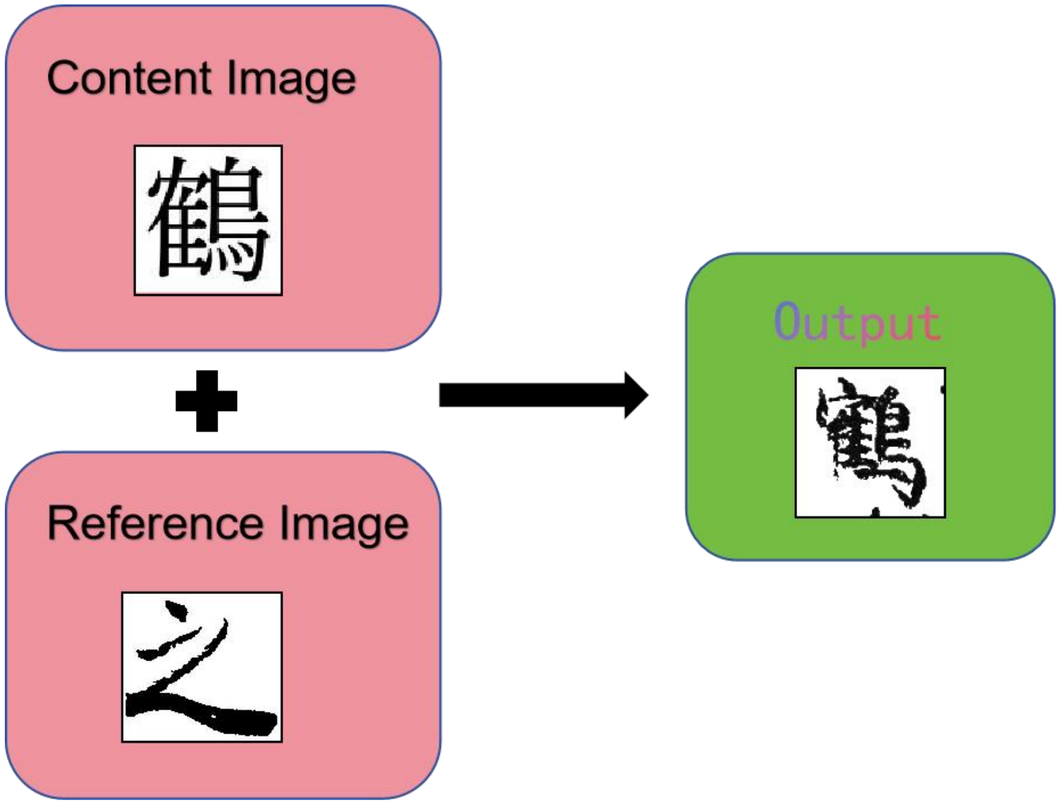
Yang, Zhenhua, et al. "Fontdiffuser: One-shot font generation via denoising diffusion with multi-scale content aggregation and style contrastive learning. AAAI, 2024

Problem

- Existing font generation methods do not perform well in generating Kuzushiji font.

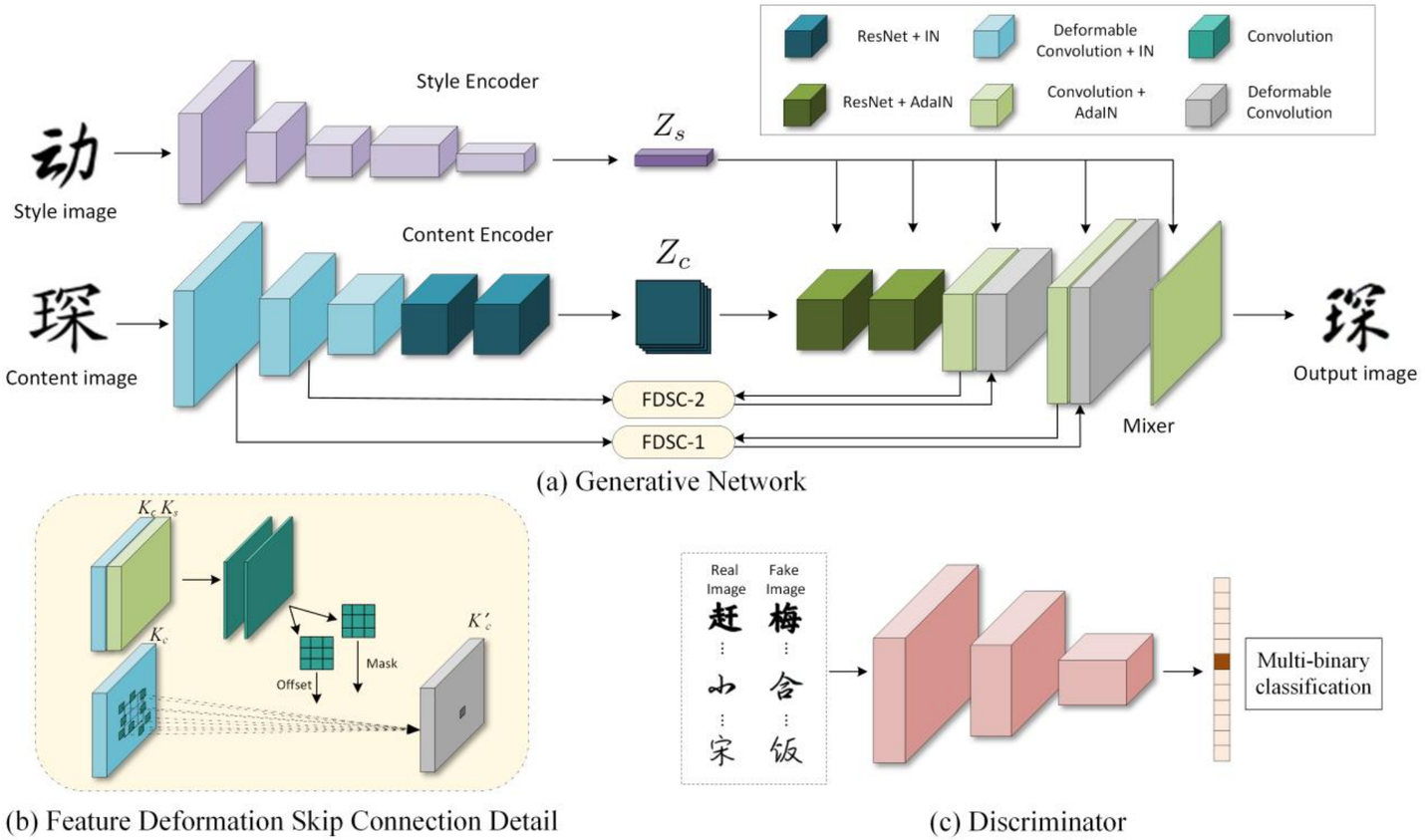
Purpose

- Transfer modern font characters to Kuzushiji characters.
- Generating high-quality Kuzushiji characters.



Related Work

Dg-font[1] achieved unsupervised learning by introducing a deformation skip connection using GAN.

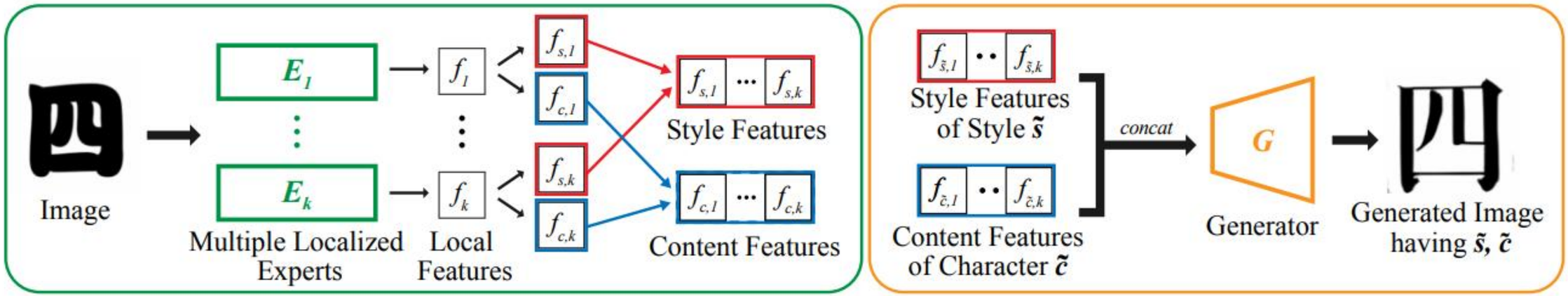


[1]Xie, Yangchen, et al. "Dg-font: Deformable generative networks for unsupervised font generation. CVPR. 2021

Related Work

- MX-Font[2] extracting multiple style features by multiple experts

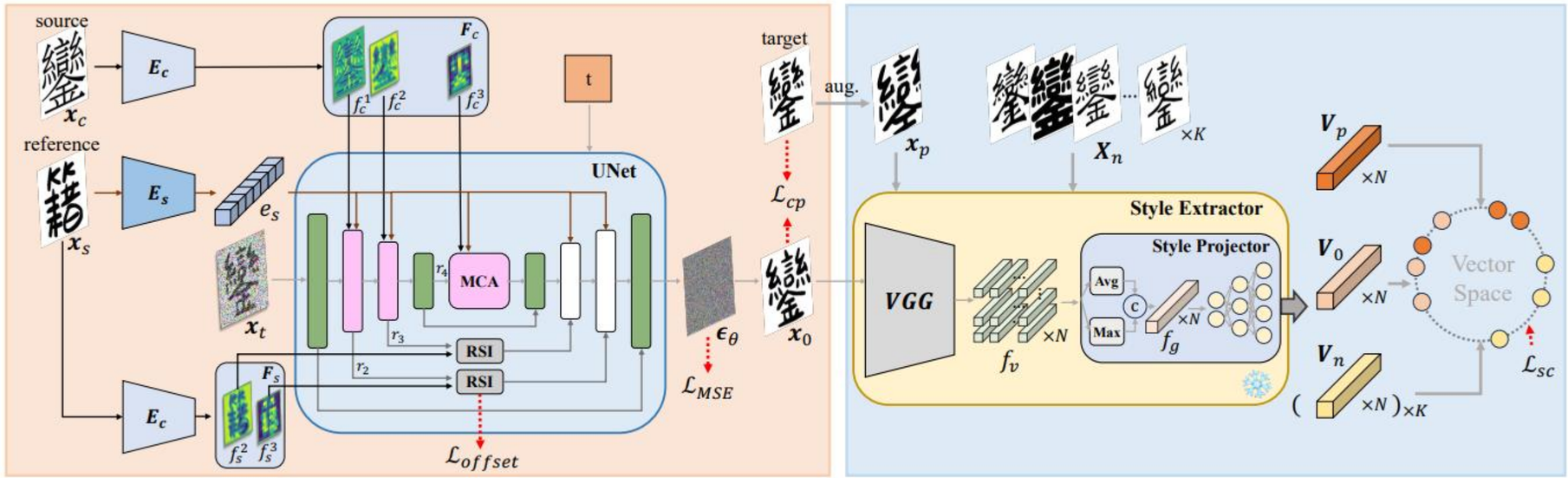
- Without being explicitly conditioned on component labels.



[2]Park, S. "Multiple heads are better than one: Few-shot font generation with multiple localized experts." ArXiv abs/2104 887 (2021).

Related Work

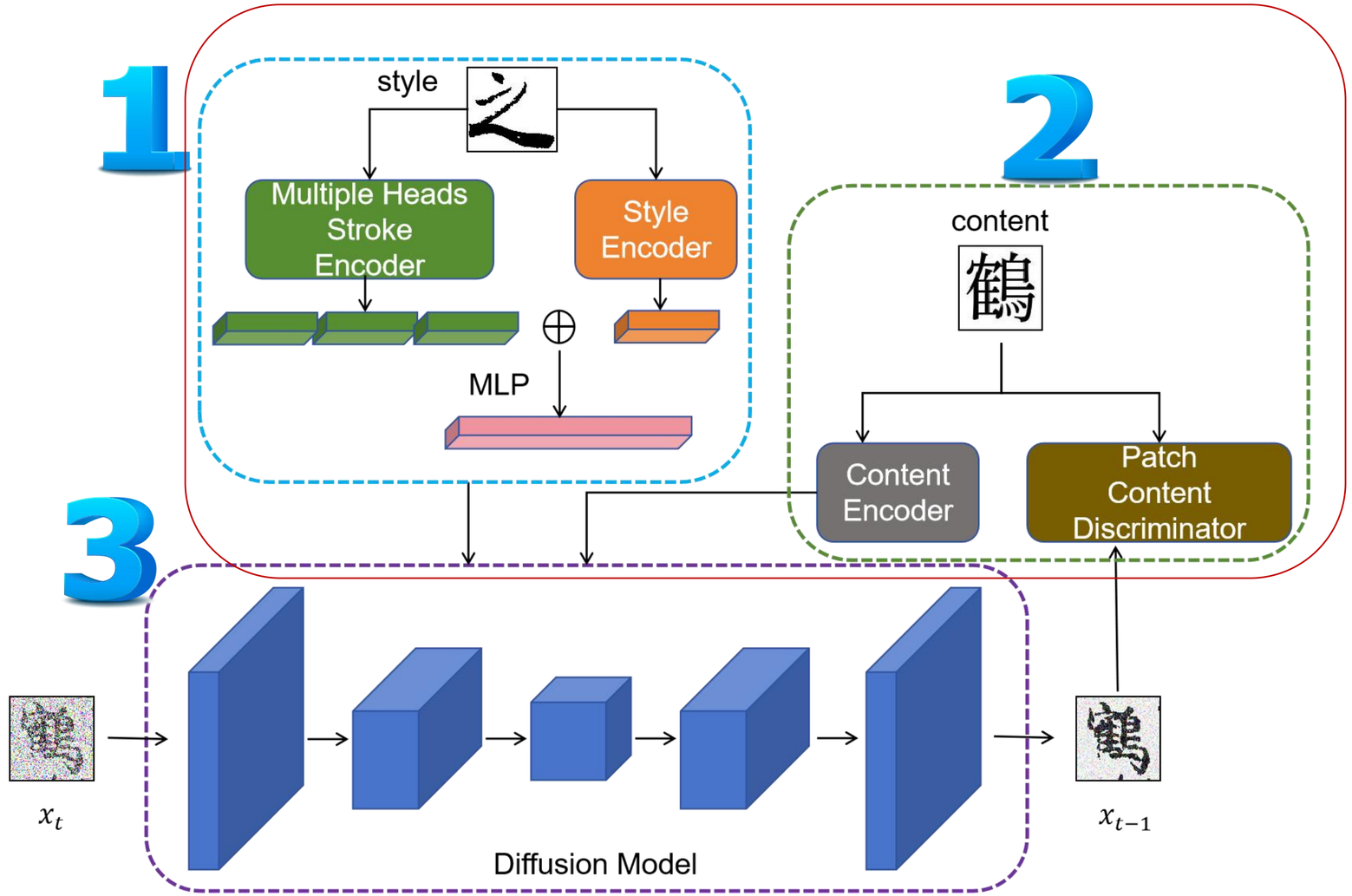
- FontDiffuser[3] leveraged multi-scale content features and an innovative style contrastive learning strategy.



[3]Yang, Zhenhua, et al. "Fontdiffuser: One-shot font generation via denoising diffusion with multi-scale content aggregation and style contrastive learning. AAAI, 2024

Proposed Method

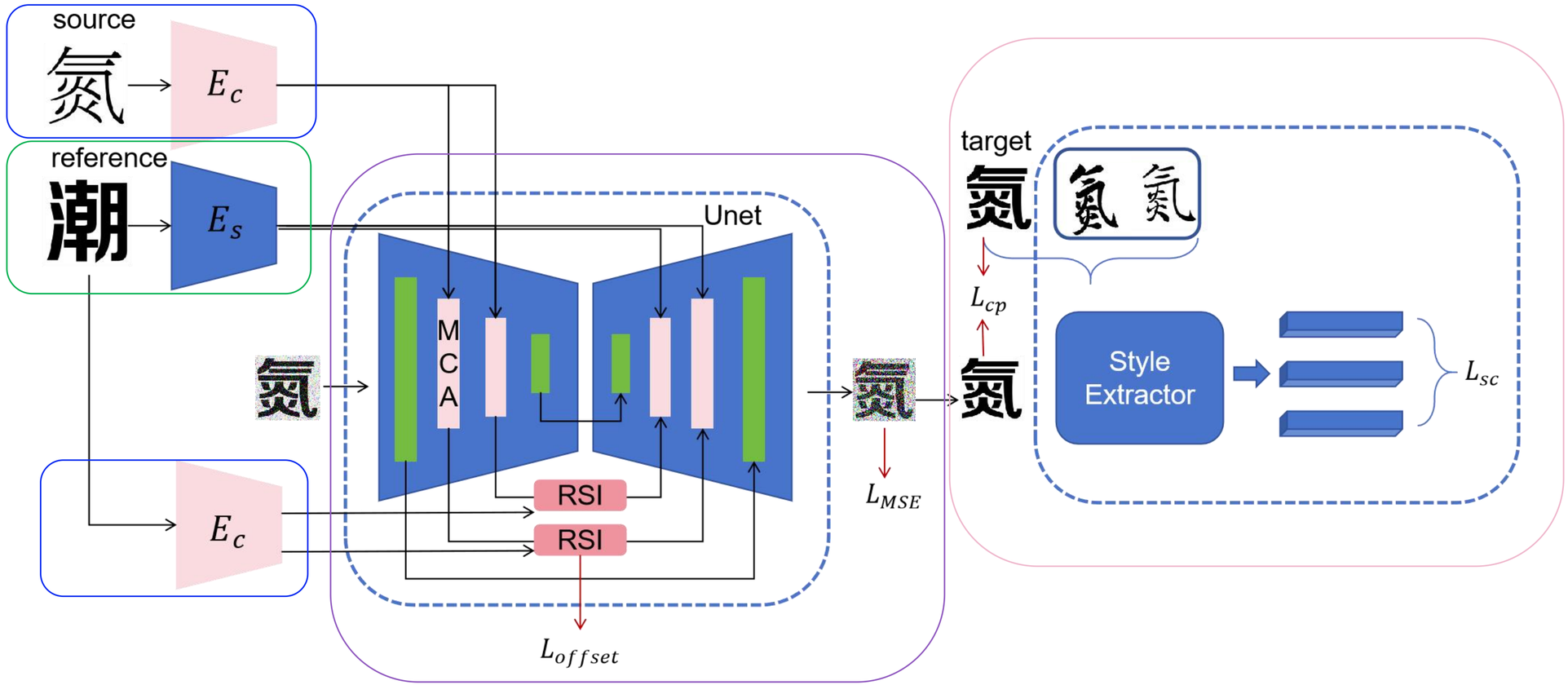
Overview of the proposed method



Proposed Method

Basic network Fontdiffuser

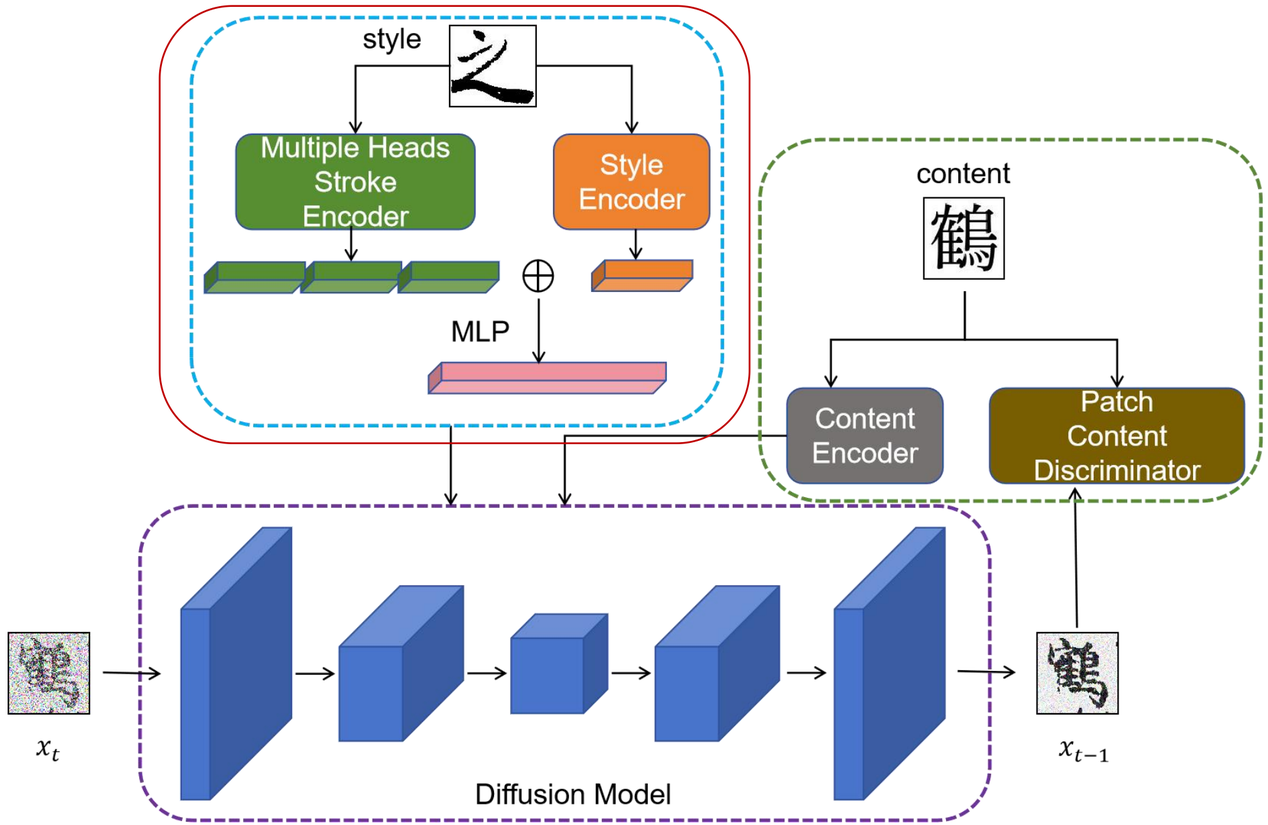
- Used as our basic diffusion model.



Proposed Method

Stroke style model

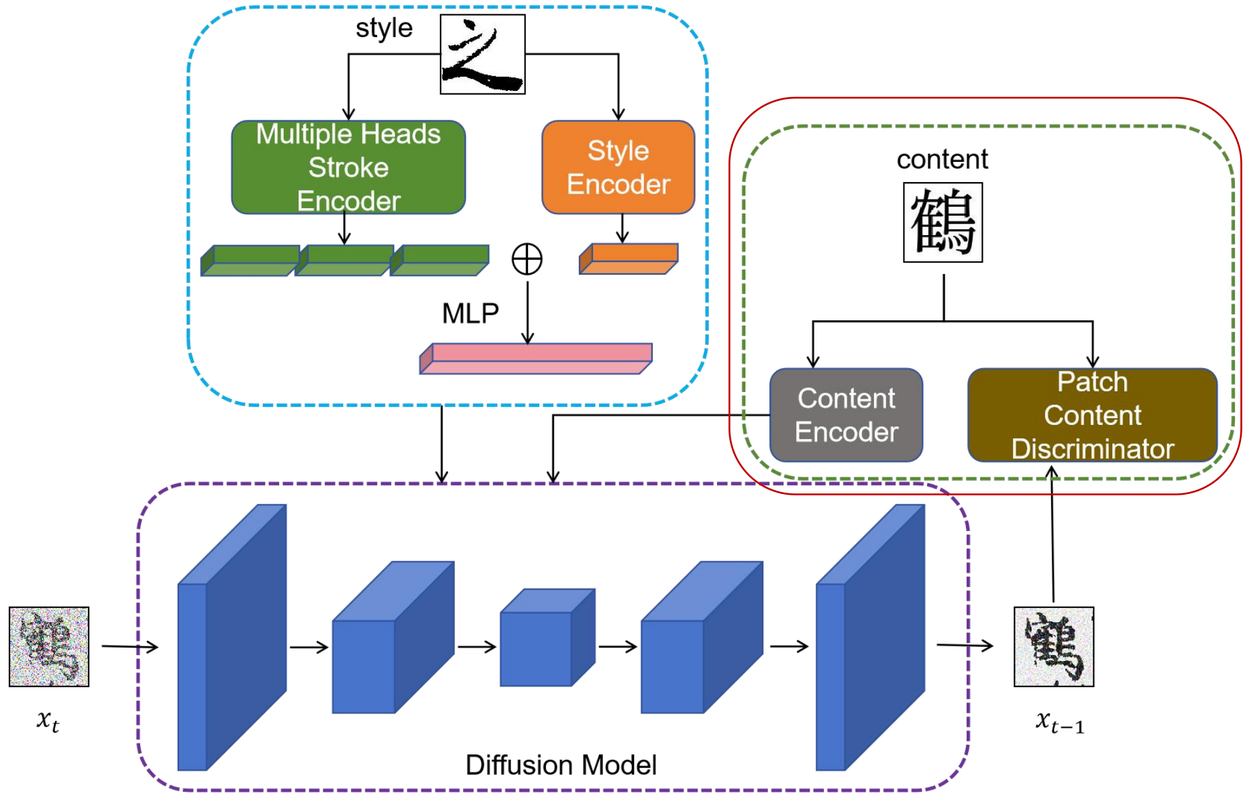
- FontDiffuser is difficult to emphasize details in strokes.
- Therefore, we propose a stroke-level feature style module.



Proposed Method

Content retaining model

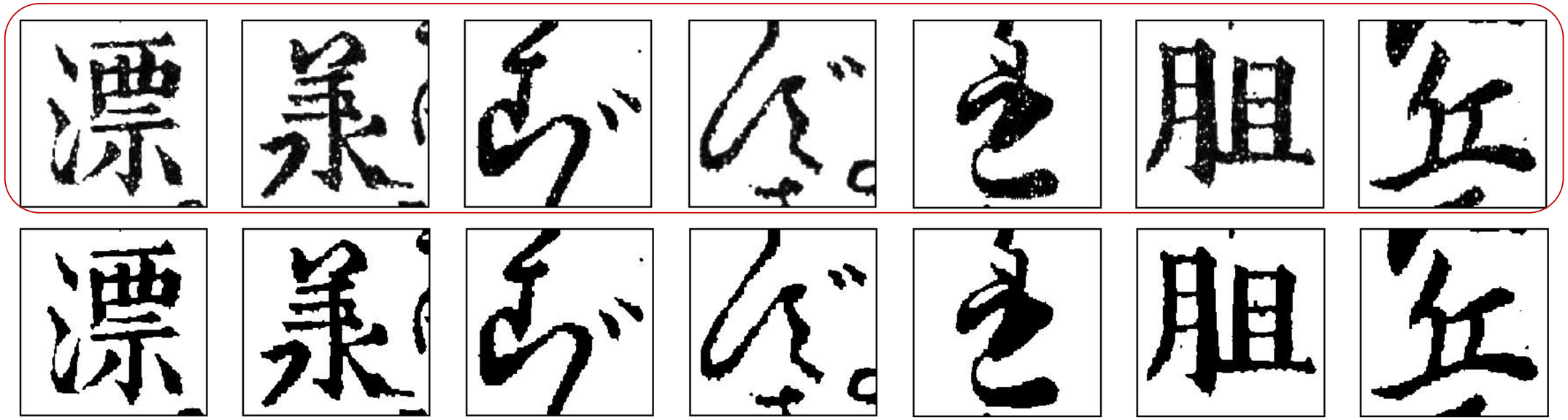
- For Reduce the instability in the font structure.
- We introduce a font content discriminator.



Data pre-processing

- Kuzushiji fonts are often of low resolution and have significant noise.
- We applied morphological operations to the images to eliminate noise in the fonts.

Original Kuzushiji data



Experiments

- We collected about 4,300 Kuzushiji font images from “Oragaharu” to train our network.

- Input : one Kuzushiji reference and content image.

- 50 inference steps.
- Our inference time takes about 7 to 10 seconds.

Experiments

- The results generated by our method.

Experiments

- Comparison results

Experiments

Quantitative evaluation results

Achieved the best and second-best results.

	FID↓	LPIPS↓	L1↓	RMSE↓	SSIM↑
FontDiffuser	1.4768	0.4844	**0.6071**	**0.5001**	<u>0.2988</u>
FontDiffuser(retrain)	<u>1.1811</u>	<u>0.4237</u>	0.6446	0.5217	0.2955
MX-Font	1.2392	0.4307	0.6459	0.5210	0.2400
CF-Font	1.1555	0.5274	0.7456	0.5639	0.2104
DG-Font	1.5265	0.5495	0.7226	0.5585	0.2006
Ours	**1.1219**	**0.3978**	<u>0.6417</u>	<u>0.5182</u>	**0.3029**
	5.01%	6.11%	(-)	(-)	1.37%

Discussion

Remaining challenges

- Image can only maintain roughly similar characteristics
  - The same strokes in different texts may be written differently.
- Not enough Kuzushiji categories can be generated.

Conclusion

- We proposed a diffusion model-based method for generating Kuzushiji fonts. It consists of :
  1. The stroke-level style extractor
  2. The content-retaining module
  3. Conditional Diffusion model

- Our method achieved state-of-the-art results.