



EUROPEAN CONFERENCE ON COMPUTER VISION

MILANO
2024

One-DM: One-Shot Diffusion Mimicker for Handwritten Text Generation

Gang Dai*, Yifan Zhang*, Quhui Ke, Qiangya Guo, Shuangping Huang

South China University of Technology, National University of Singapore,

Skywork AI, Pazhou Laboratory

TUE-AM-7598



天工
Skywork



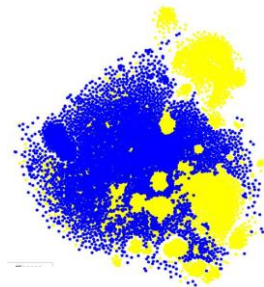
琶洲实验室
PAZHOU LAB

Overview

- 1 Background
- 2 Proposed Method
 - Style-enhanced Module
 - Gate Mechanism
 - Conditional Diffusion Model
- 3 Experimental Results
- 4 Conclusion

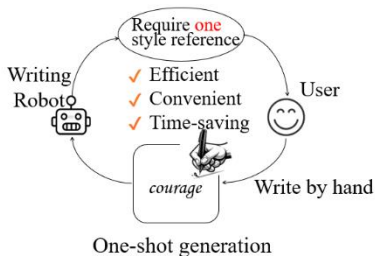
Handwritten Text Generation

- Accelerate the process of handwritten font design or potentially assist people with hand impairments
- Enrich the datasets to train efficient text recognition systems



One-shot handwritten text generation

- Our goal is to imitate user's writing style from **only a single reference**, and generate stylized handwritten texts with any content



Style sample	Textual content	Ours	DALL-E3	SD	Artbreeder
<i>quite</i>	fall				
<i>with</i>	never				
<i>when</i>	asking				

Challenges

Limitation of existing methods

- Previous handwritten text generation methods are unsatisfactory:
 - Mostly require users to **provide a few reference samples** (typically 15), making them inconvenient to use
 - Rarely achieve one-shot generation and perform poorly in emulating handwriting styles, due to their **simple style encoder design** (*e.g.*, *CNN*)

Challenges

- **Accurately extract writing style** from only one provided reference sample
- **Background noise** is commonly present in the style samples, further increasing the difficulty of style extraction

Overview

- 1 Background
- 2 Proposed Method
 - Style-enhanced Module
 - Gate Mechanism
 - Conditional Diffusion Model
- 3 Experimental Results
- 4 Conclusion

Motivation

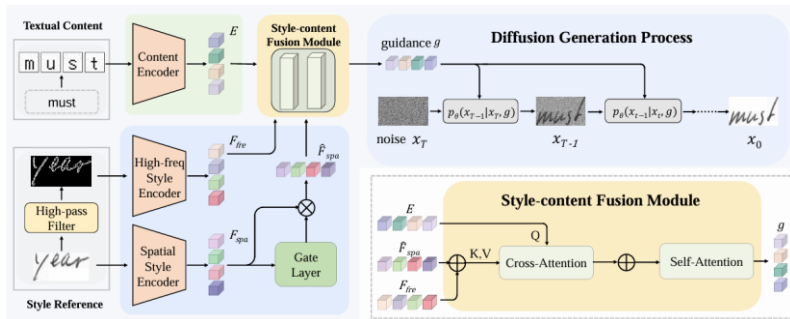


Motivation

- High-frequency components have more pronounced character contours, **clearly showcasing the style patterns** (character slant and cursive connections)
- **Incorporate the high-frequency information** of the reference sample to enhance the style extraction

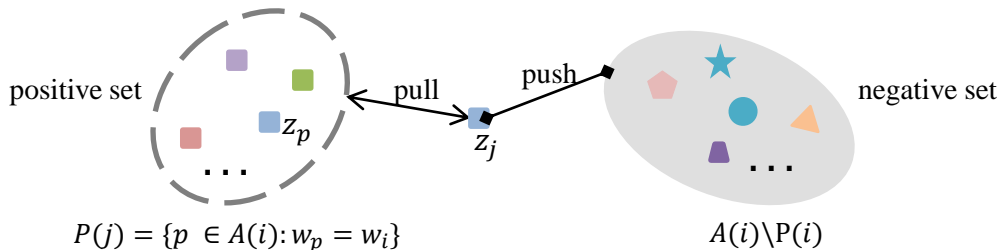
Overall Pipeline

One-DM: One-Shot Diffusion Mimicker for Handwritten Text Generation



- The style-enhanced module **independently extract spatial and high-frequency style features** from reference sample and its high-frequency information
- The gate mechanism **selectively filters out background noise** from the reference style features, allowing only meaningful style patterns to pass

One-shot Style Learning

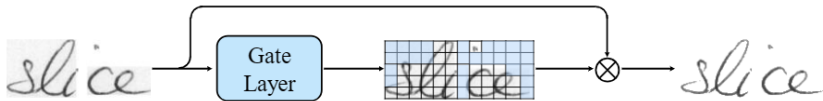


Laplacian style extraction

- Use **Laplacian kernel** to extract high-frequency components from one-shot sample
- Align the high-frequency information from the same writer to explicitly encourage to learn **discriminative writing style** (e.g., glyph slant)

$$\mathcal{L}_{Lap_NCE} = -\frac{1}{N} \sum_{i \in \mathcal{M}} \frac{1}{|P(i)|} \sum_{p \in P(i)} \log \frac{\exp(\text{sim}(e_i, e_p)/\tau)}{\sum_{a \in A(i)} \exp(\text{sim}(e_i, e_a)/\tau)}$$

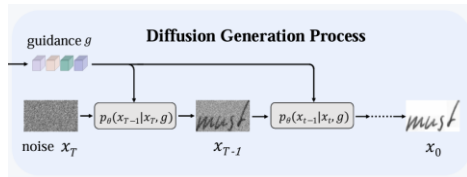
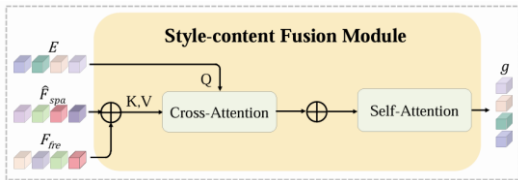
One-shot Style Learning



Background noise suppression

- Extracted spatial style features are fed into a **gate layer** to obtain the corresponding **gate units**
- Each unit determines the pass rate for the corresponding feature, enabling a **higher pass rate** for informative style features

Conditional Diffusion Model



- Learned content feature E , and two style features F_{spa} and F_{fre} , are first fused in a cross-attention, followed by a self-attention, to obtain a merged condition g
- The diffusion model progressively synthesizes the handwritten text images with **controllable content and style** conditioned on g

Overview

- 1 Background
- 2 Proposed Method
 - Style-enhanced Module
 - Gate Mechanism
 - Conditional Diffusion Model
- 3 Experimental Results
- 4 Conclusion

Handwritten text generation

■ Quantitative Comparisons with the SOTA methods

Method	Shot	Styled Evaluation				Style-agnostic	
		IV-S	IV-U	OOV-S	OOV-U	FID↓	GS↓
TS-GAN [9]	One	118.56	128.75	127.11	136.67	20.65	4.88×10^{-2}
GANwriting [27]	Few	120.07	124.30	125.87	130.68	28.37	5.67×10^{-2}
HiGAN+ [15]	One	117.33	116.95	121.55	121.48	22.95	2.06×10^{-2}
GC-DDPM [12]	One	99.86	105.73	112.52	118.39	19.05	1.31×10^{-2}
WordStylist [44]	One	98.10	104.27	109.45	115.52	18.58	2.85×10^{-2}
HWT [5]	Few	109.25	106.90	116.55	113.52	18.99	4.41×10^{-3}
VATr [46]	Few	103.75	101.73	111.64	108.76	16.03	1.74×10^{-2}
Ours (One-DM)	One	89.47	98.36	93.30	102.75	15.73	1.98×10^{-3}

- Our One-DM outperforms other methods across all evaluation metrics
 - Even surpasses the previous methods that **require 15x more references**

Handwritten text generation

■ Qualitative Comparisons with the SOTA methods

Style examples	See win next has COME Some might have the be is unassalable of air for	IN our there they on as had inside are show have us God is important	and turning which of to is you will the them may the river run from	And to is a ll Common Wood less is other in work and it than of
GANw.	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart
*HiGAN+	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart
*WordS.	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart
HWT	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart
VATr	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart
*Ours	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart	The greatest test of courage on earth is to bear defeat without losing heart

■ Our One-DM yields higher-quality results in terms of style imitation and structure preservation

Application to other languages

■ Quantitative evaluations of our One-DM and competitors

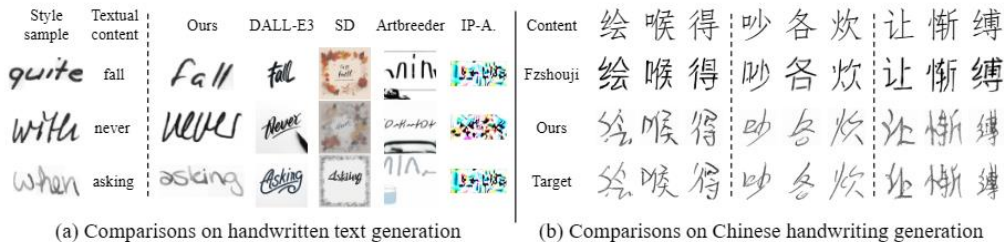


Table 5: Quantitative comparisons with competitors on styled handwritten character generation in terms of FID.

Method	Chinese	Japanese
GANwriting [27]	116.49	111.86
HWT [5]	165.74	148.66
VATr [46]	139.91	124.98
WordStylist [44]	34.61	101.93
Ours (One-DM)	27.24	95.43

- Our One-DM can handle handwritten characters in different languages and ensure the quality of synthetic samples
- Outperform other competitors in terms of FID

■ Qualitative Comparisons with SOTA industrial image generation methods



- Our One-DM excels industrial methods in style mimicry and content preservation in different languages

Overview

- 1 Background
- 2 Proposed Method
 - Style-enhanced Module
 - Gate Mechanism
 - Conditional Diffusion Model
- 3 Experimental Results
- 4 Conclusion

Conclusion

Contributions

- We propose One-DM for stylized high-quality handwritten text generation, which **only requires a single reference sample as style input**
- We introduce the high-frequency components of the reference sample to **enhance the extraction of handwriting style**

Future work

- Explore the potential of One-DM in font generation and vector font creation

The End



Code: <https://github.com/dailenson/One-DM>