

IF-Font: Ideographic Description Sequence-Following Font Generation

Xinping Chen^{1,2}, Xiao Ke^{*1,2}, Wenzhong Guo^{1,2}

¹Fujian Provincial Key Laboratory of Networking Computing and Intelligent Information Processing, College of Computer and Data Science, Fuzhou University

²Engineering Research Center of Big Data Intelligence, Ministry of Education, Fuzhou 350116, China

Overview

- 1 Introduction
 - Background
 - Motivation
- 2 Method
 - Overview of IF-Font
 - Ideograph Description Sequence
 - Fusion Module
- 3 Experiment
 - Qualitative Comparison
 - Quantitative Comparison
- 4 Conclusion
 - Conclusion

Few-shot Font Generation

骨缠悄翹



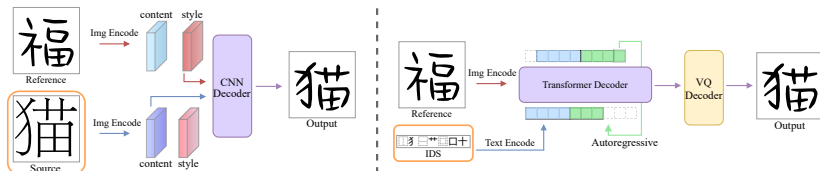
骨缠悄翹

Source

Target

Few-shot Font Generation is a technique to **simulate the target style** with just a handful of reference glyphs.

Motivation

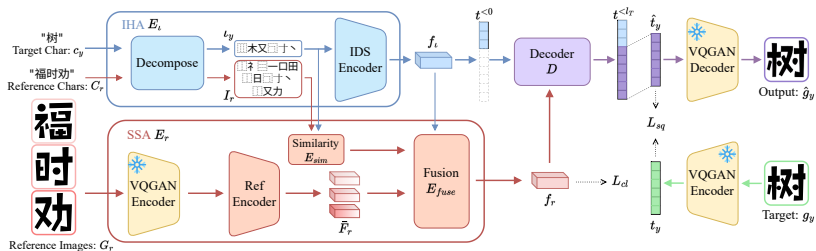


Existing Approaches (Left) assumes that a glyph can be decomposed into two distinct attributes: content and style.

IF-Font (Right) abandons the previous content-style disentanglement paradigm and generates glyphs through **next-token prediction**.

The structure of IF-Font

The overall structure of the IF-Font:



Ideograph Description Sequence (IDS)

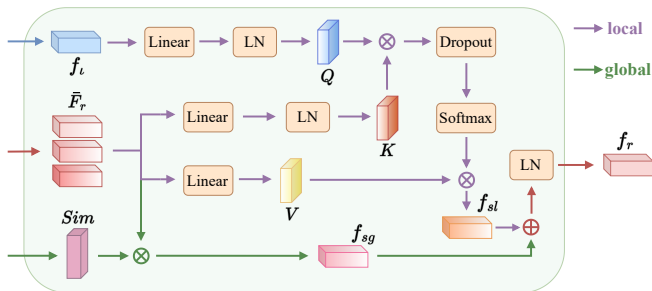
The examples of IDSs:

$$\begin{array}{l} \text{树} = \text{[]}[木][寸] = \text{[]}[木][对] \\ \text{[]}[木][又][寸] = \text{[]}[]}[木][又][寸] = \text{[]}[木][]}[又][寸] \\ \\ \text{克} = \text{[]}[古][儿] = \text{[]}[十][兄] \\ \text{[]}[十][口][儿] = \text{[]}[]}[十][口][儿] = \text{[]}[十][]}[口][儿] \end{array}$$

Ideograph Description Sequence (IDS) is a **structural description** grammar for Chinese characters define by the Unicode Standard.

Fusion Module

The overall structure of the Fusion Module:



The fusion module E_{fuse} consists of two branches: global and local style feature aggregation.

Qualitative results:

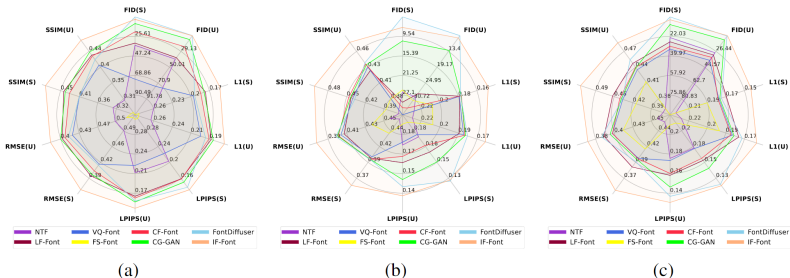
| | | | | | | | | | |
|--------------|---------|------|------|------|------|------|----|----|---|
| | Source | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| UFSC | CG-GAN | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| | LF-Font | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| | FS-Font | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| | CF-Font | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| | VQ-Font | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| | NTF | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| FontDiffuser | Source | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| Ours | Source | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| Target | Source | 骨缠梢翘 | 歉兜断葡 | 状无庸 | 绊痰 | 雷竣 | 忘旋 | 挽锋 | 海 |
| UFUC | CG-GAN | 茫固猜疯 | 婉惹呵拗 | 风曝仅敛 | 家家图步 | 扁截辽仑 | 汹 | | |
| | LF-Font | 茫固猜疯 | 婉惹呵拗 | 风曝仅敛 | 家家图步 | 扁截辽仑 | 汹 | | |
| | FS-Font | 茫固猜疯 | 婉惹呵拗 | 风曝仅敛 | 家家图步 | 扁截辽仑 | 汹 | | |
| | CF-Font | 茫固猜疯 | 婉惹呵拗 | 风曝仅敛 | 家家图步 | 扁截辽仑 | 汹 | | |
| | VQ-Font | 茫固猜疯 | 婉惹呵拗 | 风曝仅敛 | 家家图步 | 扁截辽仑 | 汹 | | |
| | NTF | 茫固猜疯 | 婉惹呵拗 | 风曝仅敛 | 家家图步 | 扁截辽仑 | 汹 | | |
| FontDiffuser | Source | 茫固猜疯 | 婉惹呵拗 | 风曝仅敛 | 家家图步 | 扁截辽仑 | 汹 | | |
| Ours | Source | 茫固猜疯 | 婉惹呵拗 | 风曝仅敛 | 家家图步 | 扁截辽仑 | 汹 | | |
| Target | Source | 茫固猜疯 | 婉惹呵拗 | 风曝仅敛 | 家家图步 | 扁截辽仑 | 汹 | | |

Quantitative results:

| Methods | | 1shot | | | | | 3shot | | | | | 8shot | | | | | User (%) \uparrow |
|----------------|-------------------|------------------|-----------------|--------------------|-------------------|-----------------|------------------|-----------------|--------------------|-------------------|-----------------|------------------|-----------------|--------------------|-------------------|-----------------|---------------------|
| | | FID \downarrow | L1 \downarrow | LPIPS \downarrow | RMSE \downarrow | SSIM \uparrow | FID \downarrow | L1 \downarrow | LPIPS \downarrow | RMSE \downarrow | SSIM \uparrow | FID \downarrow | L1 \downarrow | LPIPS \downarrow | RMSE \downarrow | SSIM \uparrow | |
| UFSC | CG-GAN [25] | 11.3911 | 0.1784 | 0.1500 | 0.3997 | 0.4428 | 10.8713 | 0.1771 | 0.1464 | 0.3982 | 0.4441 | 11.1332 | 0.1764 | 0.1457 | 0.3974 | 0.4440 | 14.78 |
| | LF-Font [38] | 32.9264 | 0.1764 | 0.1586 | 0.3967 | 0.4465 | 29.1840 | 0.1786 | 0.1576 | 0.3998 | 0.4432 | 26.9590 | 0.1694 | 0.1567 | 0.3875 | 0.4590 | 11.42 |
| | FS-Font [47] | 112.0971 | 0.2836 | 0.3108 | 0.5145 | 0.2795 | 25.5231 | 0.2075 | 0.1916 | 0.4343 | 0.3865 | 93.7912 | 0.1900 | 0.2086 | 0.4124 | 0.4183 | 6.55 |
| | CF-Font [50] | 20.4457 | 0.1839 | 0.1581 | 0.4066 | 0.4323 | 30.8426 | 0.1767 | 0.1650 | 0.3977 | 0.4468 | 30.9829 | 0.1784 | 0.1595 | 0.3990 | 0.4465 | 12.22 |
| | VQ-Font [36] | 72.7064 | 0.1958 | 0.2215 | 0.4201 | 0.4077 | 32.9390 | 0.1789 | 0.1775 | 0.4016 | 0.4405 | 33.6378 | 0.1774 | 0.1732 | 0.3995 | 0.4413 | 7.50 |
| | NTF [10] | 35.3797 | 0.2602 | 0.2027 | 0.4887 | 0.3244 | 26.1215 | 0.2275 | 0.1749 | 0.4542 | 0.3659 | 23.0519 | 0.2238 | 0.1739 | 0.4501 | 0.3720 | 6.43 |
| | FontDiffuser [55] | 3.9969 | 0.1938 | 0.1371 | 0.4180 | 0.4076 | 3.6989 | 0.1774 | 0.1248 | 0.3980 | 0.4370 | 4.1017 | 0.1748 | 0.1234 | 0.3947 | 0.4420 | 18.32 |
| IF-Font (Ours) | <u>6.7695</u> | 0.1529 | 0.1307 | 0.3688 | 0.4915 | <u>6.8359</u> | 0.1478 | <u>0.1258</u> | 0.3620 | 0.5021 | <u>6.7383</u> | 0.1429 | 0.1216 | 0.3552 | 0.5140 | 22.78 | |
| UFUC | CG-GAN [25] | 13.4734 | 0.1805 | 0.1508 | 0.4019 | 0.4362 | 13.0347 | 0.1790 | 0.1471 | 0.4001 | 0.4383 | 13.2049 | 0.1780 | 0.1462 | 0.3991 | 0.4391 | 15.73 |
| | LF-Font [38] | 37.3840 | 0.1835 | 0.1620 | 0.4047 | 0.4283 | 28.8252 | 0.1850 | 0.1609 | 0.4071 | 0.4283 | 30.5147 | 0.1735 | 0.1582 | 0.3920 | 0.4473 | 11.65 |
| | FS-Font [47] | 112.6636 | 0.2847 | 0.3112 | 0.5155 | 0.2764 | 31.2833 | 0.2106 | 0.1923 | 0.4373 | 0.3785 | 98.9486 | 0.1921 | 0.2095 | 0.4146 | 0.4131 | 6.17 |
| | CF-Font [50] | 22.8601 | 0.1865 | 0.1584 | 0.4094 | 0.4259 | 34.0245 | 0.1796 | 0.1660 | 0.4009 | 0.4399 | 33.2477 | 0.1809 | 0.1601 | 0.4019 | 0.4399 | 12.03 |
| | VQ-Font [36] | 75.1737 | 0.1980 | 0.2217 | 0.4223 | 0.4018 | 36.4831 | 0.1809 | 0.1776 | 0.4037 | 0.4345 | 36.5486 | 0.1796 | 0.1733 | 0.4017 | 0.4354 | 8.23 |
| | NTF [10] | 39.3581 | 0.2678 | 0.2074 | 0.4958 | 0.3086 | 29.9205 | 0.2303 | 0.1753 | 0.4568 | 0.3593 | 27.9580 | 0.2290 | 0.1755 | 0.4553 | 0.3619 | 6.33 |
| | FontDiffuser [55] | 8.2524 | 0.1914 | 0.1527 | 0.4157 | 0.4163 | 7.6444 | 0.1771 | 0.1413 | 0.3981 | 0.4418 | 8.9166 | 0.1702 | 0.1367 | 0.3890 | 0.4543 | 18.57 |
| IF-Font (Ours) | <u>8.4844</u> | 0.1651 | 0.1387 | 0.3845 | 0.4676 | <u>8.4922</u> | 0.1597 | 0.1338 | 0.3775 | 0.4782 | <u>8.3203</u> | 0.1561 | 0.1305 | 0.3728 | 0.4864 | 21.28 | |

Our IF-Font achieves **SOTA** with different numbers of reference glyphs.

Quantitative results



IF-Font achieves **SOTA** on all metrics under three few-shot settings. The metrics are annotated with brackets in the figure to specify the dataset used for evaluation: (S) represents UFSC and (U) refers to UFUC. (a) 1-shot setting. (b) 3-shot setting. (c) 8-shot setting.

Conclusion

Main contributions:

- We propose IF-Font, which abandons the previous content-style disentanglement paradigm and generates glyphs through **next-token prediction**.
- We devise a novel **IDS Hierarchical Analysis (IHA) module** that analyzes the spatial structures and components of Chinese characters. It allows our decoder flexibly control the generated content with the encoded semantic features.
- Leveraging corresponding IDSs, we design the **Structure-Style Aggregation (SSA) module** to extract and efficiently aggregate the style features of reference glyphs.

The End

Thanks for watching