GC-Font: Few-Shot Font Generation via Global Contextual Feature Modelling

Weiran Chen, Guiqian Zhu, Ying Li, Yi Ji, Chunping Liu

wrchen2023@stu.suda.edu.cn, {gqzhu, ingli, jiyi, cpliu}@suda.edu.cn

School of Computer Science and Technology, Soochow University, Suzhou, China

Github Page

Abstract

Few-shot font generation aims to create new fonts with a limited number of glyph references. It can be used to greatly reduce the labour cost of manual font design. However, due to the variety and complexity of font styles, the results generated by existing methods often suffer from visible defects, such as stroke errors, blurriness or distorted shapes. To address these issues, we propose GC-Font, a novel framework which integrates a Global Contextual Feature Modelling (GCFM) module. Specifically, this module is inserted between the content encoder and decoder, where it fuses convolution and attention mechanisms to process intermediate feature maps and injects enhanced global contextual features into the decoder. Moreover, we apply adaptive convolutions to the low-level feature maps from the content encoder to strengthen contextual correlations. In addition, a skeleton consistency loss and an edge consistency loss are also designed to improve geometric alignment. Extensive experiments reveal that our GC-Font outperforms the state-of-the-art methods in both qualitative and quantitative evaluations, demonstrating its effectiveness on diverse font styles. Our source code can be found at https://github.com/wrchen2001/GC-Font.

Contributions

- We devise an inventive Global Contextual Feature Modelling (GCFM) module, which consists of convolution operations and an attention mechanism to better augment the global contextual features. Based on this, we introduce GC-Font, a new font generation framework that facilitates accurate style transfer while preserving glyph integrity.
- We embed adaptive convolutions within the content encoder, enabling dynamic adjustments of convolutional filters according to the input data. Besides, we also design two auxiliary loss functions, skeleton consistency loss and edge consistency loss, to better improve the overall visual consistency.
- Extensive experimental results verify that our GC-Font surpasses the leading models in both qualitative and quantitative evaluations, which demonstrates its effectiveness and generalizability on diverse font styles and characters.

The Proposed Method

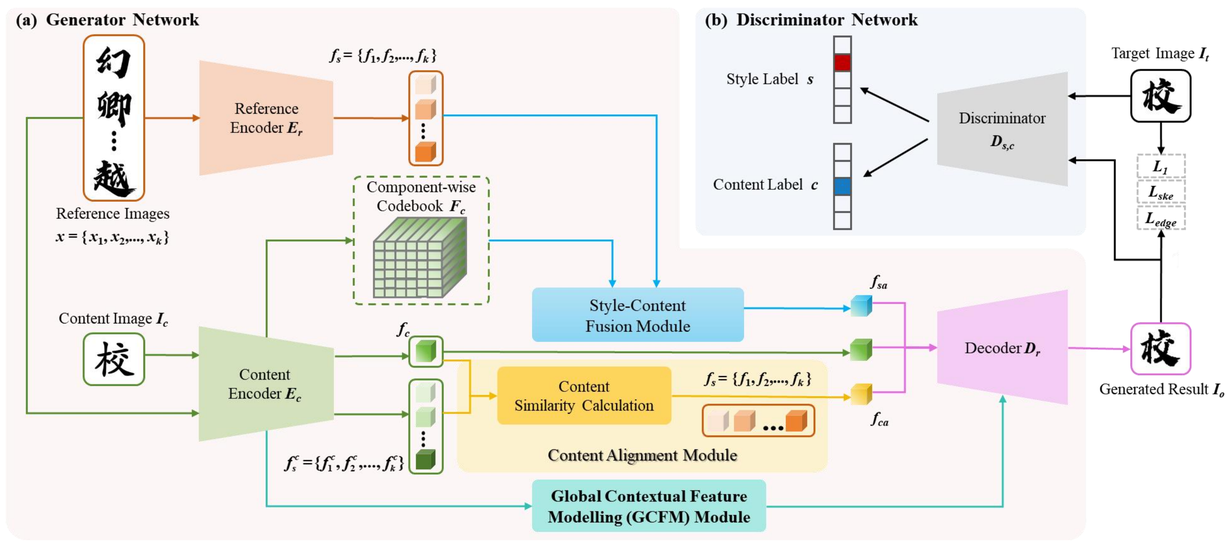
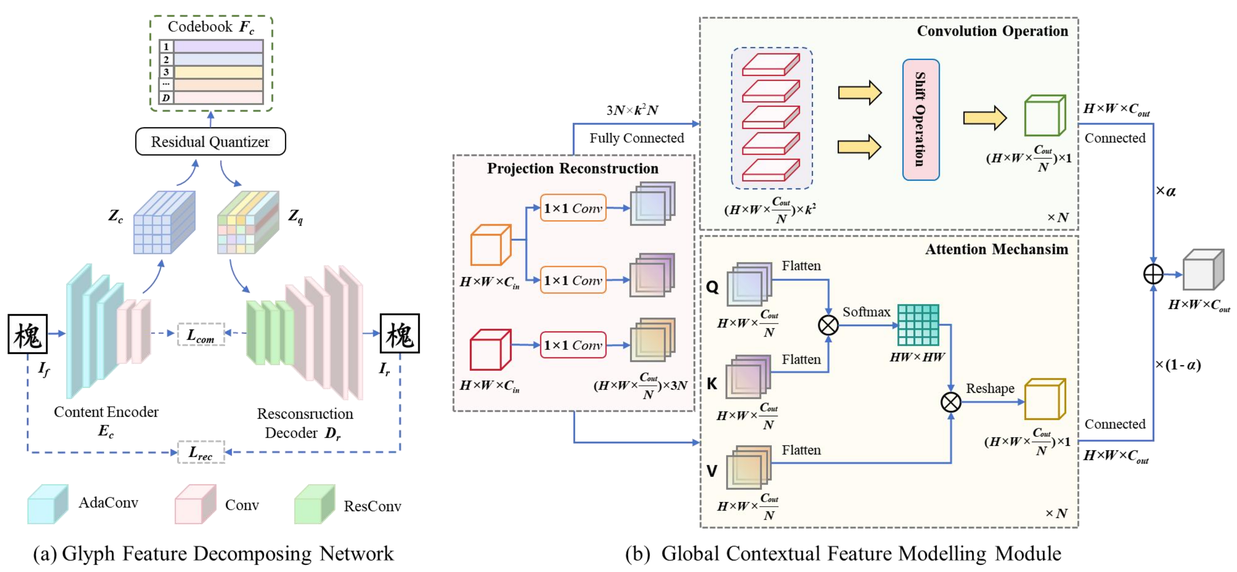


Figure 1: The framework of our proposed GC-Font. (a) The generator network consists of six key parts: a pre-trained content encoder E_c, a reference encoder E_r, a style-content fusion module, a content alignment module, a global contextual feature modelling module and a decoder D_r. (b) The discriminator network $D_{s,c}$ is employed to distinguish the real and fake images, while also classifying the content and style categories of the generated characters.



(a) Glyph Feature Decomposing Network (b) Global Contextual Feature Modelling Module

Figure 2: The architecture of glyph feature decomposing network and global contextual feature modelling module in GC-Font model.

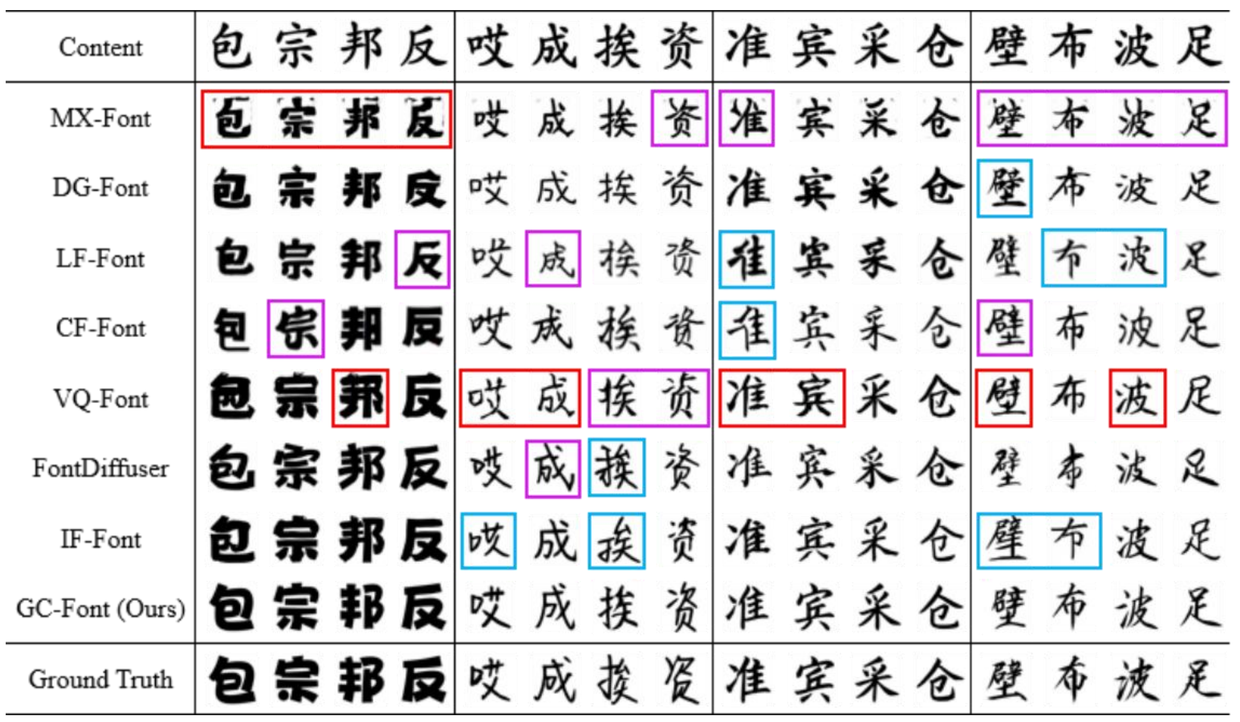
Experimental Results

➢ Quantitative comparison results

Dataset	Method	SSIM↑	RMSE↓	LPIPS↓	FID↓	L1↓	User study↑
UFUC	MX-Font	0.6966	0.3159	0.2359	62.7500	0.1235	9.55%
	DG-Font	0.6527	0.3238	0.2058	61.6022	0.1241	9.85%
	LF-Font	0.6768	0.3110	0.2516	66.8840	0.1190	9.70%
	CF-Font	0.6613	0.3102	0.2014	59.7644	0.1169	11.20%
	FontDiffuser	0.6390	0.3579	0.2816	79.4458	0.1530	10.65%
	IF-Font	0.6891	0.3130	0.2173	59.6292	0.1038	14.50%
	GC-Font (Ours)	**0.7230**	**0.2748**	**0.1711**	**56.6042**	**0.0921**	**21.40%**
SFUC	MX-Font	0.6733	0.3314	0.2197	61.0583	0.1321	8.90%
	DG-Font	0.6478	0.3278	0.2006	58.3525	0.1277	9.80%
	LF-Font	0.6140	0.3335	0.2726	69.2239	0.1520	9.35%
	CF-Font	0.6569	0.3186	0.1974	57.9868	0.1202	11.65%
	VQ-Font	0.6414	0.3304	0.2069	56.7299	0.1305	12.35%
	FontDiffuser	0.6317	0.3691	0.2910	73.3052	0.1574	10.95%
	IF-Font	0.6652	0.3280	0.2221	62.8246	0.1213	14.20%
	GC-Font (Ours)	0.7166	0.2816	0.1781	54.5017	0.1142	22.80%

➢ Qualitative comparison results

(a) Comparison results on UFUC dataset



(b) Comparison results on UFSC dataset

Figure 3: Qualitative comparison results on UFUC and SFUC datasets. Characters in blue boxes suffer from stroke errors (including missing or redundant strokes), the red boxes indicate noticeable blurring, and the purple boxes denote distorted glyph shapes. Zoom in to see details.

BMVC 2025