Few-Shot Font Generation with Deep Metric Learning

Haruka Aoki, Koki Tsubota, Hikaru Ikuta and Kiyoharu Aizawa (The University of Tokyo)

Introduction

Few-shot font generation...

- Generate images of **all characters from a few reference samples**
- Outputs should have
  consistent style with the samples
  content information (what the letter is)

A few samples → Whole character set (≒Font)

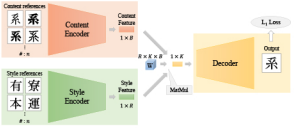
A G J ➡ A B C D E F G H I J ⋯

永 中 右 ➡ 理 有 台 塗 章 本 系 和 ⋯

Related Works

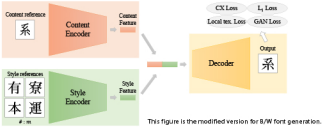
1. EMD [Zhang+, CVPR18]

The architecture for B/W Chinese font



2. AGIS-Net [Gao+, TOG19]

The architecture with GAN

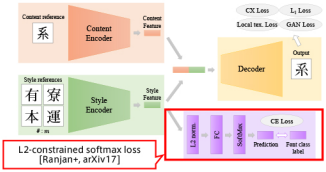


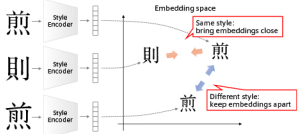
This figure is the modified version for B/W font generation.

Problem: Low quality when the number of style reference is extremely limited

Proposed Method

Introducing **metric learning** to the style encoder





Style encoder can focus only on **style** characteristics, independent from the content of samples

Experiment

Backbone networks: EMD, AGIS-Net
# of style reference samples: n ∈ {5, 10, 15, 30}

Dataset: Japanese typographic fonts

- 368 fonts (338 for training, 30 for evaluation)
- 2965 glyphs for each font
- 64 x 64 grayscale image

	n = 5		n = 10		n = 15		n = 30	
	L1	SSIM	L1	SSIM	L1	SSIM	L1	SSIM
AGIS-Net	0.306	0.554	0.275	0.607	0.245	0.640	0.208	0.684
AGIS-Net + DML (ours)	**0.274**	**0.566**	**0.219**	**0.656**	**0.202**	**0.686**	**0.193**	**0.705**
EMD	0.189	0.701	**0.179**	0.721	0.176	0.727	0.174	0.733
EMD + DML (ours)	**0.187**	**0.705**	**0.179**	**0.726**	**0.174**	**0.730**	**0.170**	**0.738**

dim. of embeddings was 512; parameter m for AGIS-Net was 4.

We observed remarkable improvement, especially when n is extremely limited