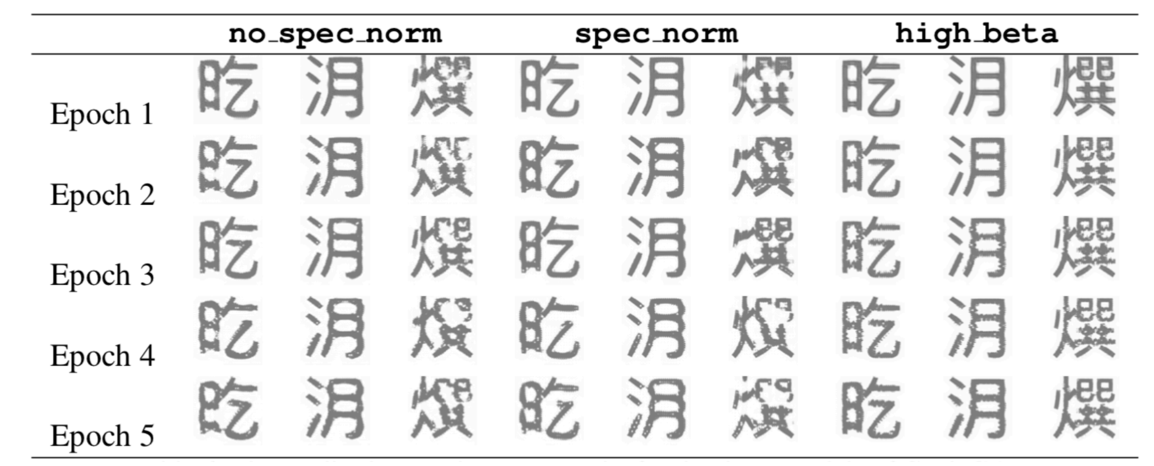
Cross-Lingual Font Style Transfer

Stanford University

Peiyu Liao
Department of Computer Science
pyliao@stanford.edu

Fang-Yu Lin
Department of Electrical Engineering
fangyuln@stanford.edu



Problem

Designing a new Chinese font style is costly and inefficient since a standard font file contains nearly fifty thousand characters.

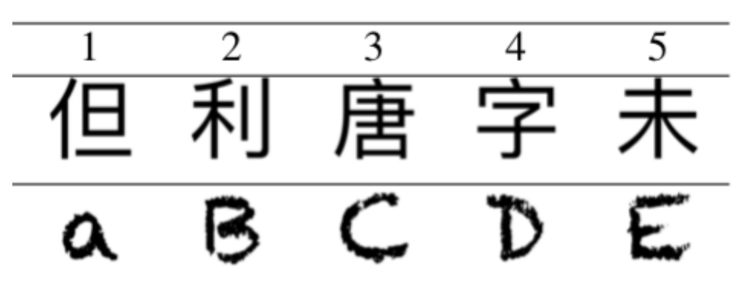
In this project, a cross-lingual font style transfer model is built to apply existing English font styles to Chinese characters by image style transfer. This way, an existing font design of any language can potentially be applied to any other language characters in the world.

Dataset

English Font Image Dataset
Chalkduster, 512 English fonts

Traditional Chinese Font Image Dataset
PingFang, 48900 Chinese fonts

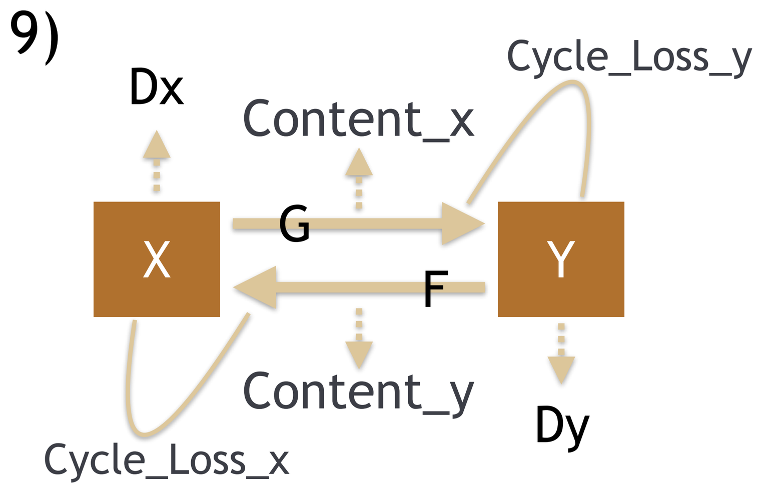
Exported as 128 × 128 3 RGB images



Data Splitting	Content Training	Style Training	Content Testing	Style Testing
NST	N/A	N/A	256	1
ITN	48644	1	256	1
CycleGAN	48644	512	256	1

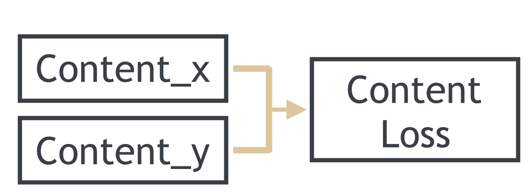
Method

Baseline
Neural Style Transfer: Loss Network(VGG19)
Image Transformation Network:
Transform Network + Loss Network(VGG19)

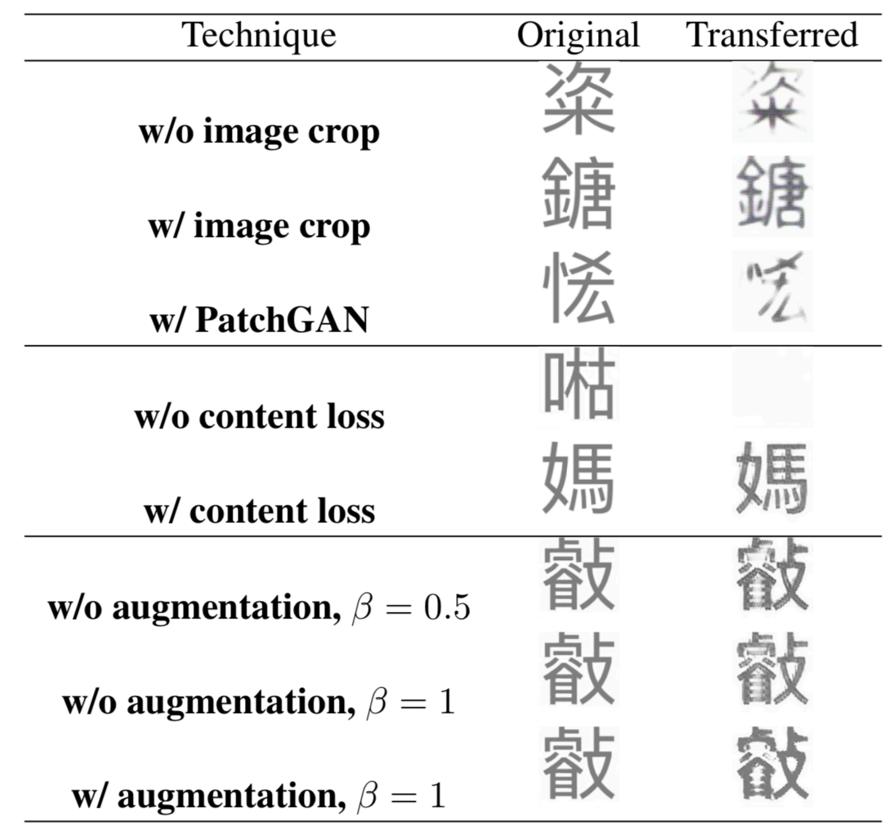


CycleGAN
Image Cropping / PatchGAN
Content Loss
Data Augmentation
Content Loss Weight

$$\begin{aligned}&\mathcal{L}(G, F, D_X, D_Y)\\ =&\mathcal{L}_{\text{GAN-G}}(G, D_Y) + \mathcal{L}_{\text{GAN-D}}(G, D_Y)\\ +&\mathcal{L}_{\text{GAN-G}}(F, D_X) + \mathcal{L}_{\text{GAN-D}}(F, D_X)\\ +&\lambda\mathcal{L}_{\text{cyc}}(G, F) + \alpha\mathcal{L}_{\text{identity}}(G, F) + \beta\mathcal{L}_{\text{content}}(G, F)\end{aligned}$$



CycleGAN Experiments



Problem: White-out effect

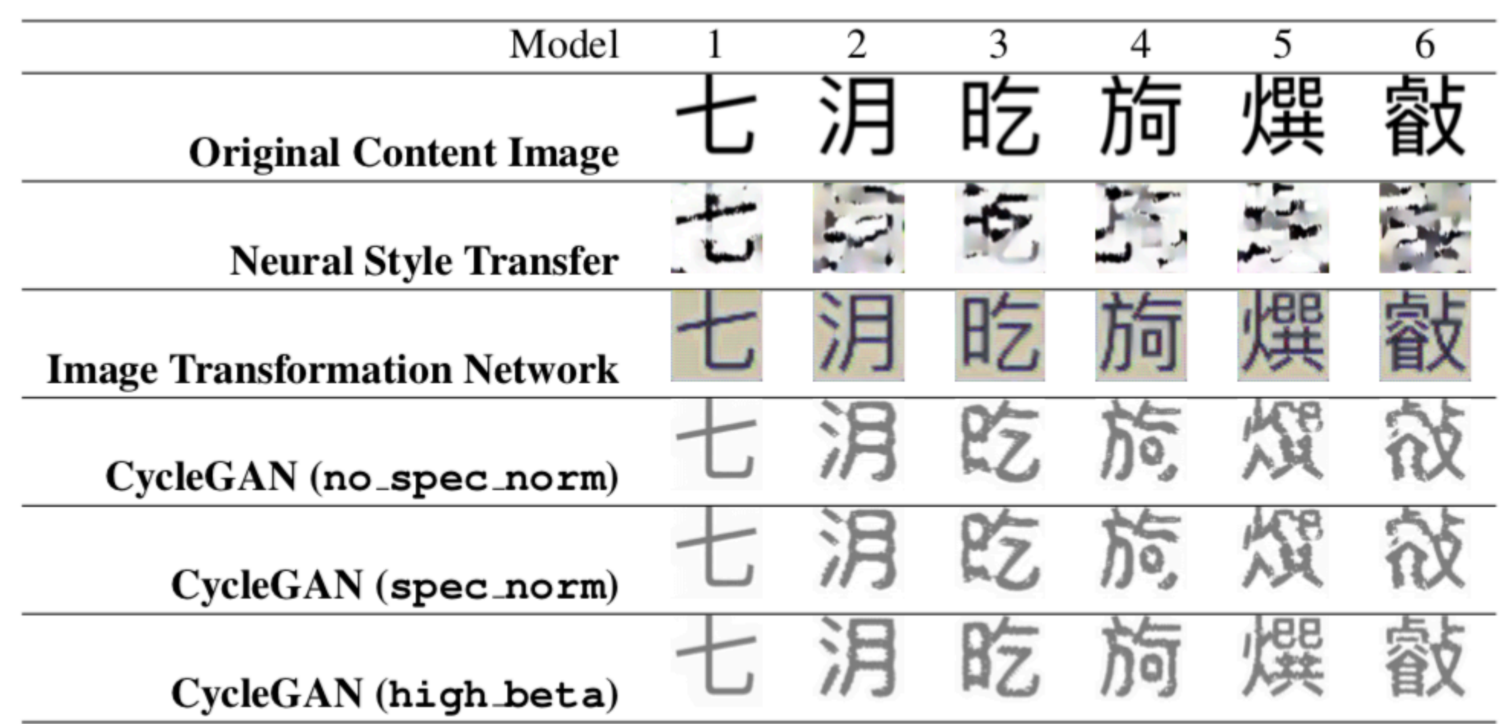
Image Cropping / PatchGAN
D learns to differentiate # of dark pixels and center-concentrated or not
Image cropping alleviated most of the white-out effect

Content Loss
D might weight other features more than the color information.
G and D learned penalization from content loss, shift to tuning the stroke style.

Content Loss Weight
Low, affects stroke completeness; High, learning slows down, train for more epochs.

Data Augmentation
Prevent D from memorizing style strokes from the center image crops
Providing generalization for model training

3 Model Experiments



	OCR Acc	Style Loss	Human Content Score	Human Style Score
NST	15.09%	2.5043	0.6	2.3
ITN	4.40%	30.1335	5.0	0
CycleGAN (No Spec Norm)	13.21%	2.0773	4.6	4.8
CycleGAN (Spec Norm)	20.13%	2.4889	4.4	4.8
CycleGAN (High Beta)	28.93%	2.3138	5.0	4.6

Baseline
NST can only be applied on simple structure. It preserved the general outline, but can hardly be correctly recognized. ITN performs the opposite. Content is preserved, but the style is not transferred.

CycleGAN
With higher stress of content loss in `high_beta`, the strokes maintain completeness, model learning is conservative and the style improves slowly.
`no_spec_norm` is the most aggressive in learning style. Spectral normalization might serve as a stabilizing factor.

Conclusion

NST is able to produce style pattern, but content will be distorted. ITN only reproduce images, fails to transfer the style. CycleGAN performs the best, producing recognizable characters with styling. Image cropping and content loss are the most salient techniques to prevent wiped-out.

Reference

[1] J.-Y.Zhu,T.Park,P.Isola,andA.A.Efros.Unpairedimage- to-image translation using cycle-consistent adversarial net-works. In Proceedings of the IEEE international conference on computer vision, pages 2223–2232, 2017.