D17

SceneTextStyler: Editing Text with Style Transformation

Honghui Yuan, Keiji Yanai

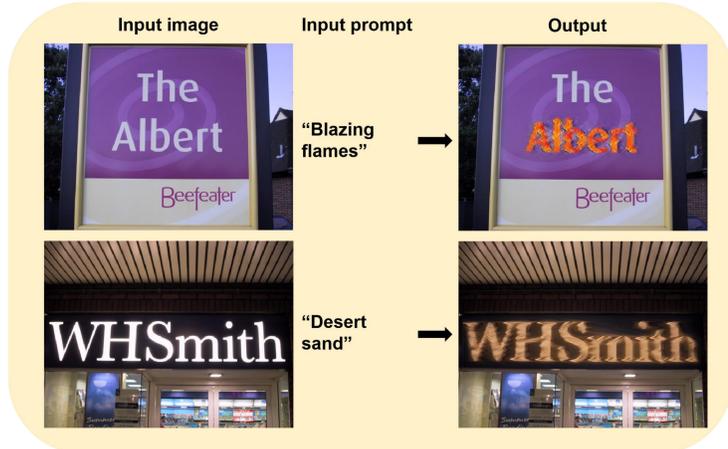The University of Electro-Communication, Tokyo, JAPAN

UEC

Abstract

We proposed a new framework named FontCLIPstyler[1] to realize **scene text style transformation.**
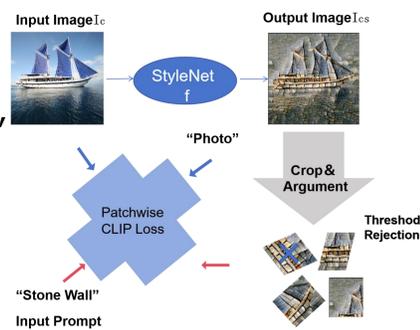
Our method could freely change the style of text areas in scene images **using prompts** based on CLIPStyler[2].
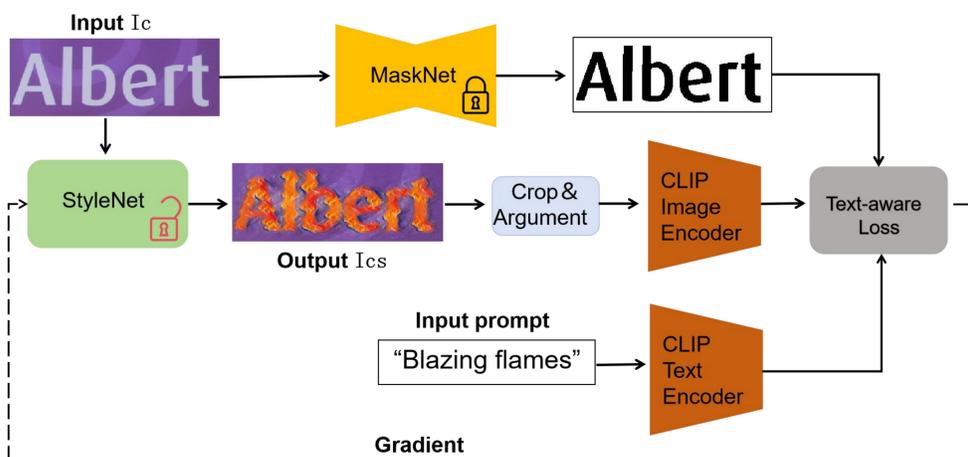
Input image   Input prompt   Output

"Blazing flames"

"Desert sand"

Methodology

Previous Work: ClipStyler[2]

- For general image style transfer, and **not applicable to text images**.
- **Unable to transform styles to specific areas** in the image.

Input Image I_c Output Image I_{cs}

StyleNet f

"Photo"

Crop & Argument

Patchwise CLIP Loss

Threshold Rejection

"Stone Wall"

Input Prompt

Proposed Method

Input I_c

MaskNet

Albert

StyleNet

Output I_{cs}

Crop & Argument

CLIP Image Encoder

Text-aware Loss

Input prompt

"Blazing flames"

CLIP Text Encoder

Gradient

1. Network for extracting text masks in images(**MaskNet**)
   - Pre-trained with 2000 images using Unet

2. Transfer semantic style of prompts(**StyleNet**)
   - Optimized with loss function Text-aware Loss

3. Only transform the style of the text area without changing the background(**Text-aware Loss**)

Text-aware Loss

$$L_{ta} = \lambda_d L_{distance} + \lambda_p L_{patch} + \lambda_r L_{recon}$$

$L_{distance}$ Allows style transformation within a limited region based on distance transformation of the input image.

L_{patch} Transfer semantic style of prompts to text area

L_{recon} Background reconstruction

[1] Honghui Yuan and Keiji Yanai: Font Style Translation in Scene Text Images with CLIPstyler. In: ICPR. 2024

[2] Kwon, G., Ye, J.C.: Clipstyler: Image style transfer with a single text condition. In: CVPR. 2022

[3] Kamra, C.G., Mastan, I.D., Gupta, D.: Sem-cs: Semantic clipstyler for text-based image style transfer. In: ICIP. 2023
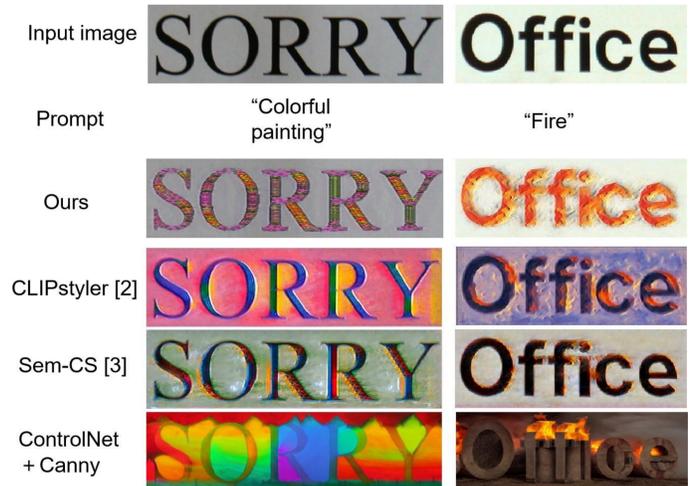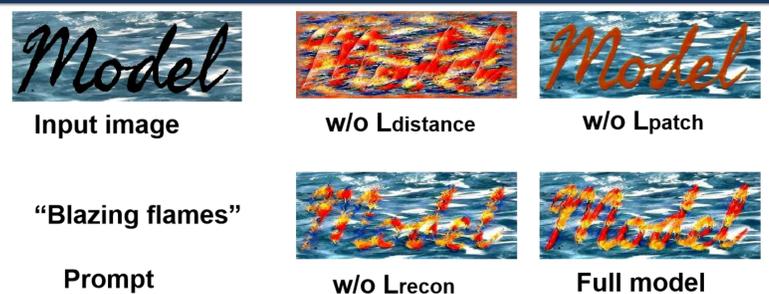
Results

1. Generation Results

Input scene images, Input prompts and Generated results

DESK — Starry Night by Vinvent van gogh

The — Fire

FAIRY — Stone wall

lock — Sunflowers

NHS — Mosaic

HELP — Green crystals

EXIT — Rocks

The — Pencil

BEER — Sunflowers

lock — Fire

DESK — Graffiti style painting

ASDA — Flower

DENTAL — The great wave off kanagawa by Hokusai

Service — Pop art of night city

Campus — Fire

2. Comparison with existing methods

Input image	SORRY	Office
Prompt	"Colorful painting"	"Fire"
Ours	SORRY	Office
CLIPstyler [2]	SORRY	Office
Sem-CS [3]	SORRY	Office
ControlNet + Canny	SORRY	Office

3. Ablation Studies

Input image   w/o L$_{distance}$ w/o L$_{patch}$

"Blazing flames" Prompt   w/o L$_{recon}$ Full model

4. Demo System

Scene Text Style Transfer

Our real-time application allows users to freely select the text in image and using prompt to generate the style scene text.

Prompt   Input   Output   Prompt   Input   Output

Blazing flames + The Albert

Desert sand + WHSmith

Oil painting of sunflowers + FAIRY

Rocks + EXIT

Style Prompt

Fire

Image

Number of text lines   0

Selected image

Example: Prompt Samples

Examples

Sunflowers   Stone wall   pop art of night city   oil painting of flowers   neon light   mosaic

graffiti style painting   Starry Night by Vincent van Gogh   The great wave off kanagawa   Acrylic painting

A fauvism style painting   Sketch in black pencil   Desert sand   Pencil   A painting of vivid flames

Image

Run Generation