Controllable Artistic Text Style Transfer via Shape-Matching GAN

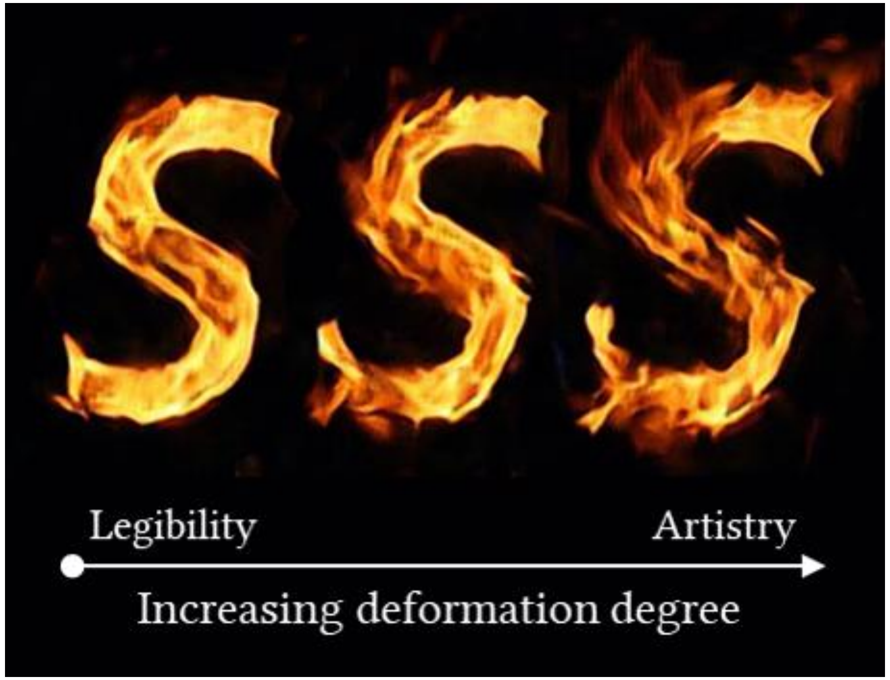
Shuai Yang[1,2], Zhangyang Wang[2], Zhaowen Wang[3], Ning Xu[3], Jiaying Liu[1] and Zongming Guo[1]

1 Institute of Computer Science and Technology, Peking University 2 Texas A&M University 3 Adobe Research

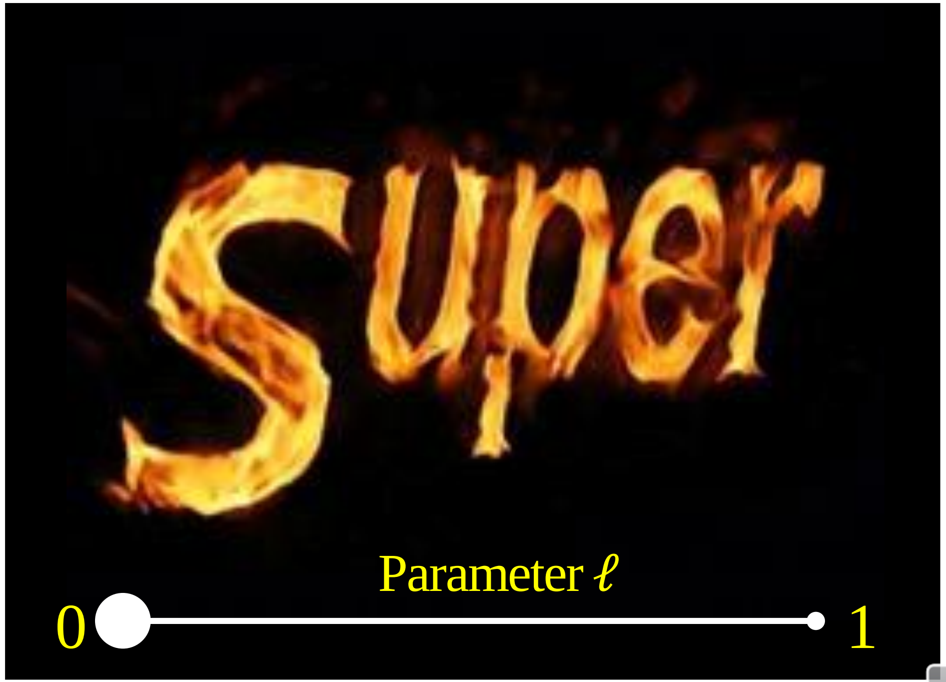
Problem: Controllable Text Style Transfer

- Input: style image, target text, deformation degree ℓ
- Output: artistic text
- Large $\ell \rightarrow$ more **artistry**; less **legibility**: balance?

Input



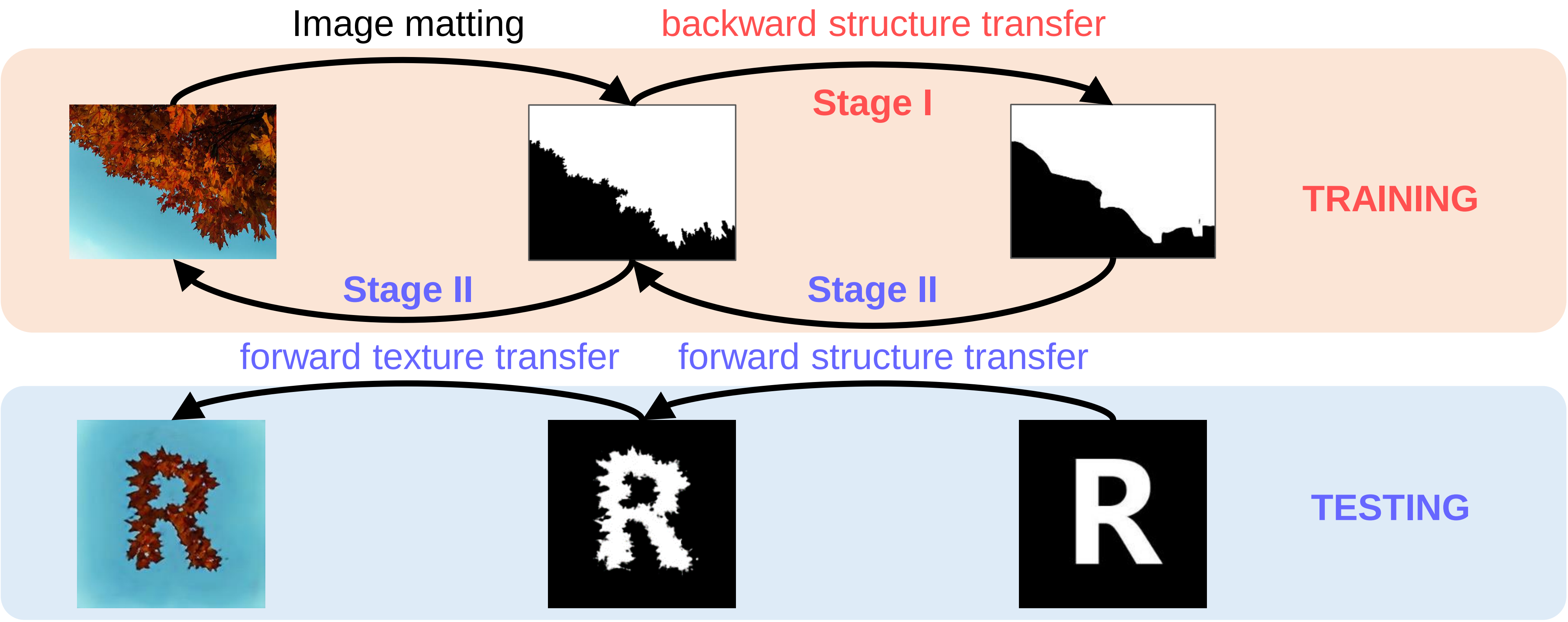
Adjust the stylistic degree of glyph



Controllable output

Problem: Controllable Text Style Transfer

- Bidirectional shape matching
  - Backward structure transfer: prepare training data
  - Forward structure transfer: learn shape deformation

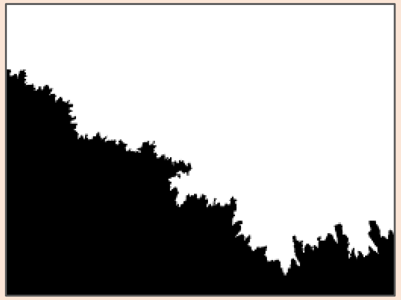


Problem: Controllable Text Style Transfer

- Challenge
  - Limited Data: one style image to train the network?
  - Controllable: one network for fast forward multiple scales

CHALLENGE I

Limited Data

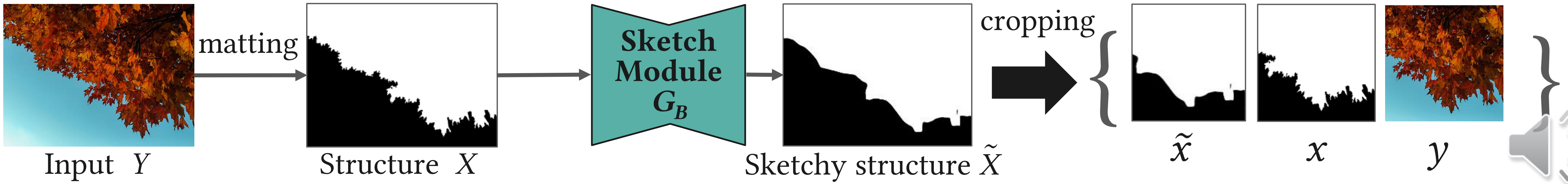


CHALLENGE II

Fast Multi-Scale Transfer

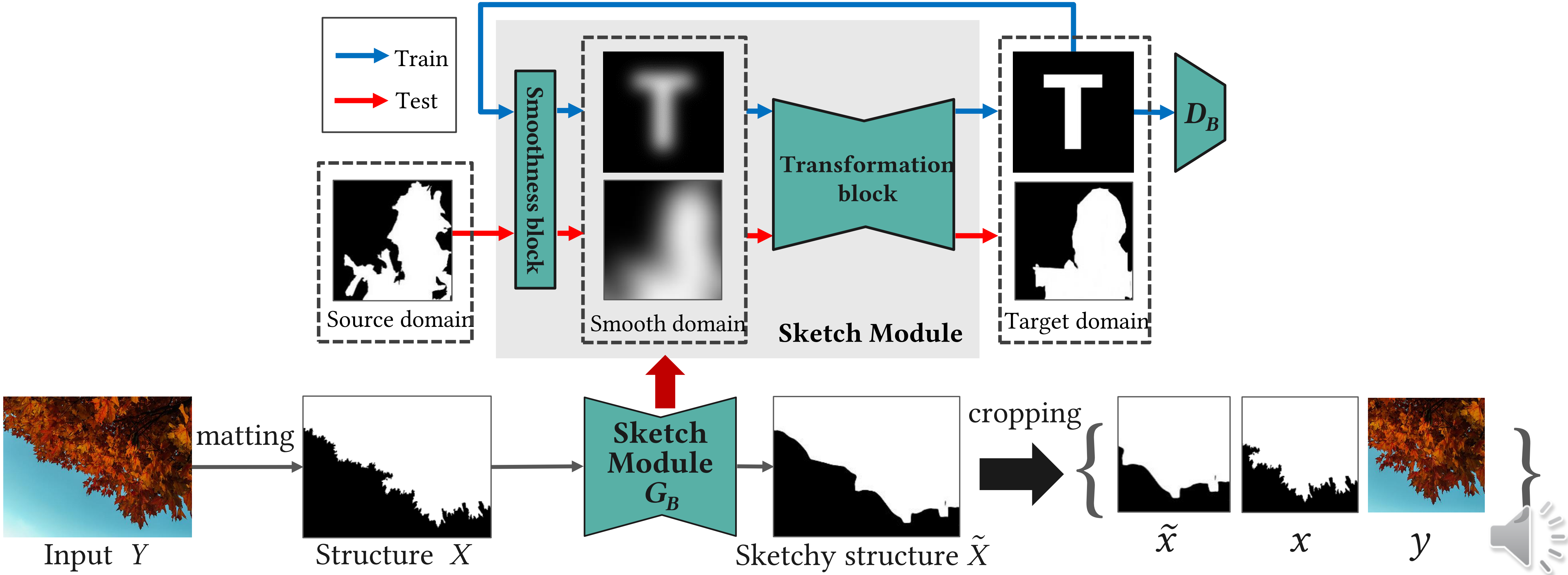
Framework

- Stage I: Input preprocessing (Backward Structure Transfer)
  - Structure map of Y: Photoshop or image matting
  - Train Sketch Module to obtain a sketchy version of X
  - **CHALLENGE I:** Limited Data
    - Generate training data: random cropping $\tilde{X}$, X, Y



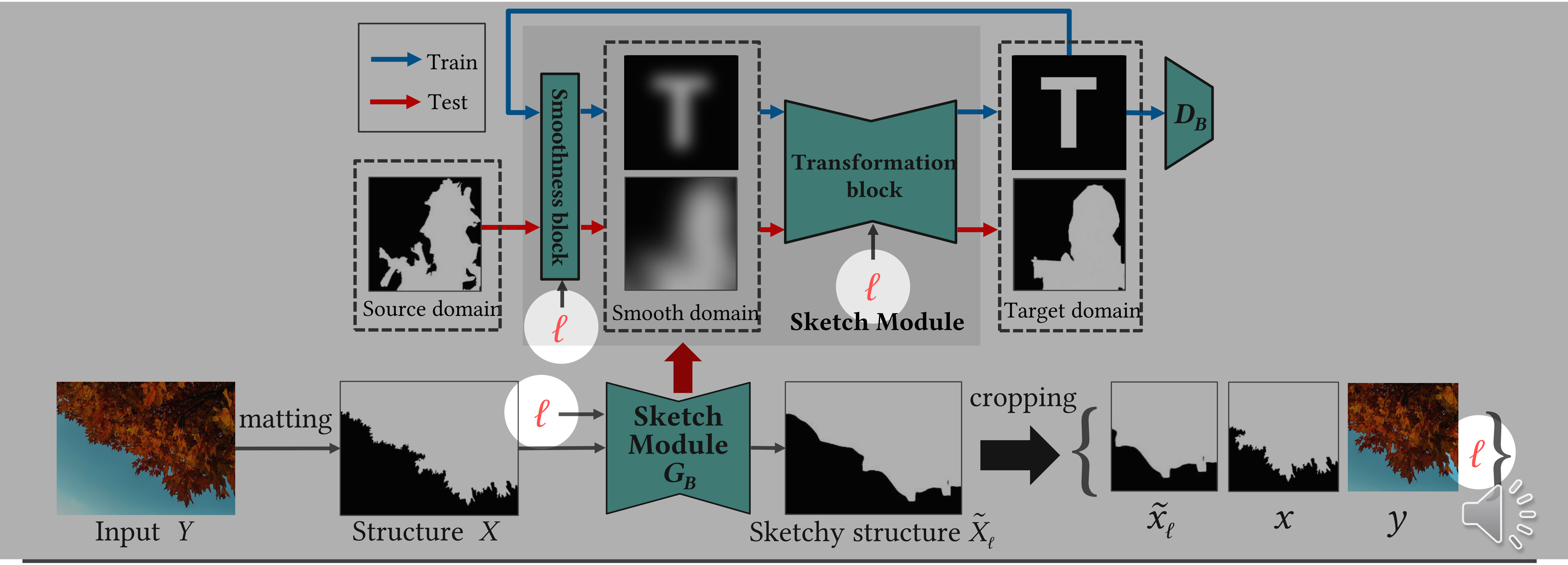
Backward Structure Transfer (G_B)

- Gaussian blur to maps T and X into a smooth domain
- Train CNN to map the smoothed image back to the text domain



Backward Structure Transfer (G_B)

- **CHALLENGE II:** Fast Multi-Scale Transfer
  - The standard deviation of Gaussian kernel is controlled by ℓ



Backward Structure Transfer (G_B)

- **CHALLENGE II:** Fast Multi-Scale Transfer
  - Multi-scale training data generation
  - More blurry → More sketchy → Higher deformation degree

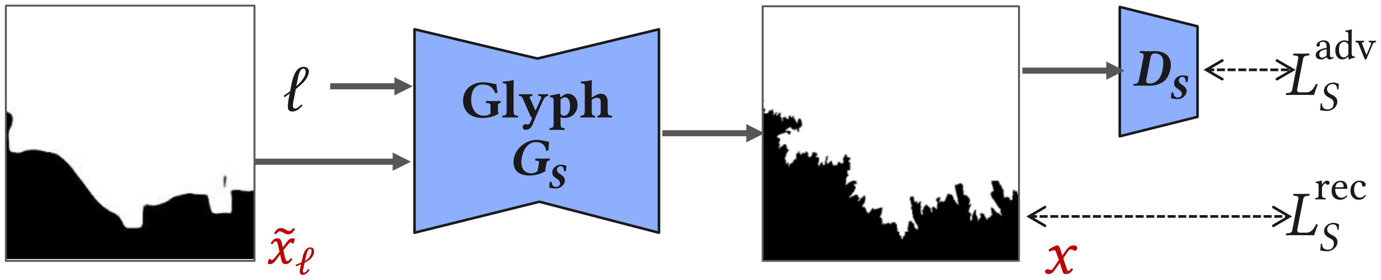
Style image

Structure map

Sketchy Structure map

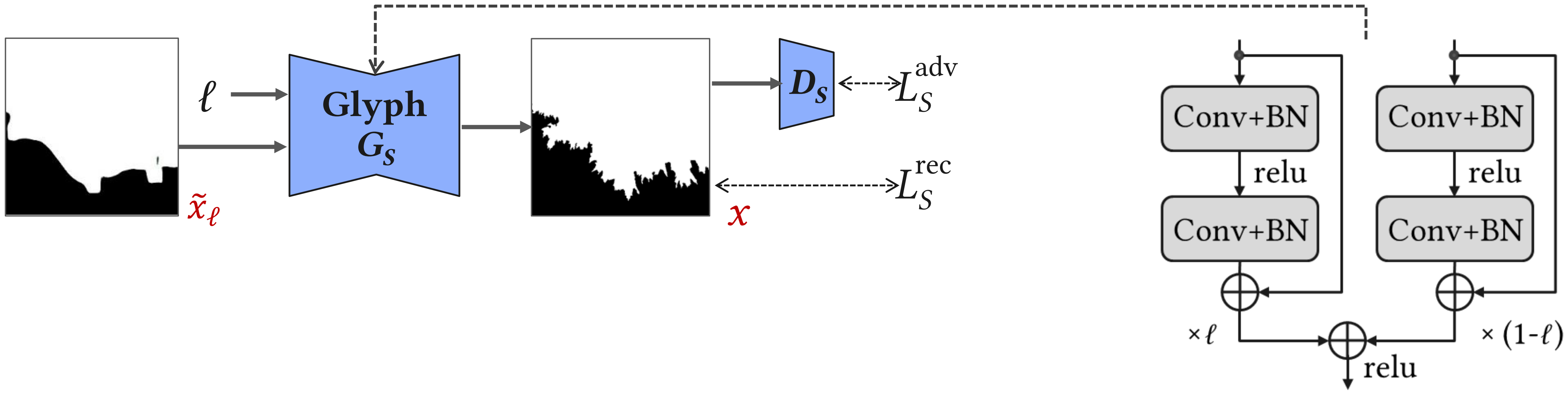
Framework

- **Stage II: Forward Structure Transfer**
  - Conditional image-to-image translation framework
  - Training: learn to map $\tilde{x}_\ell$ with different deformation degrees back to x



Forward Structure Transfer (G_S)

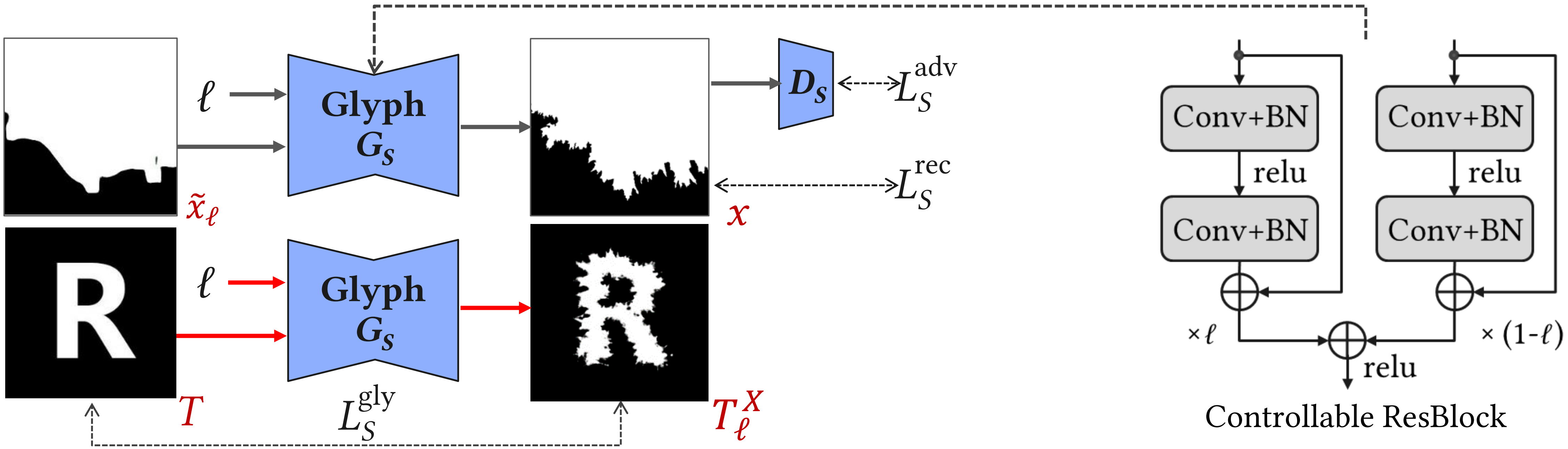
- **CHALLENGE II:** Fast Multi-Scale Transfer
  - Controllable Resblock: linear combination of 2 ResBlocks weighted by ℓ
  - ℓ = 0/1: solely deal with greatest / tiniest structure deformation
  - $\ell \in$ (0,1): compromise between the two extremes



Controllable ResBlock

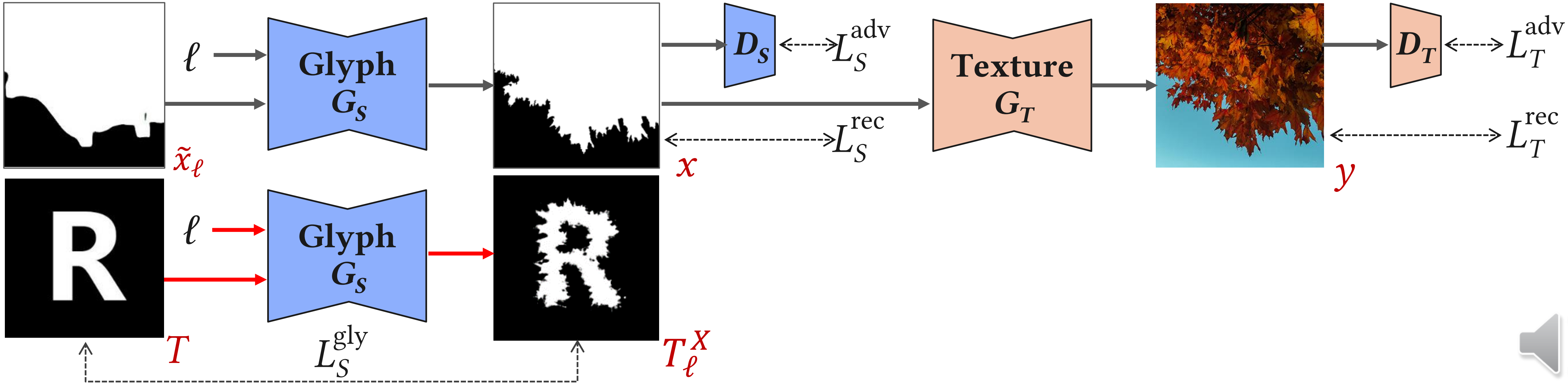
Forward Structure Transfer (G_S)

- Stage II: Forward Structure Transfer
  - Conditional image-to-image translation framework
  - Test: transfer the shape style of x onto T, producing T_ℓ^X
  - Glyph loss: text legibility preservation



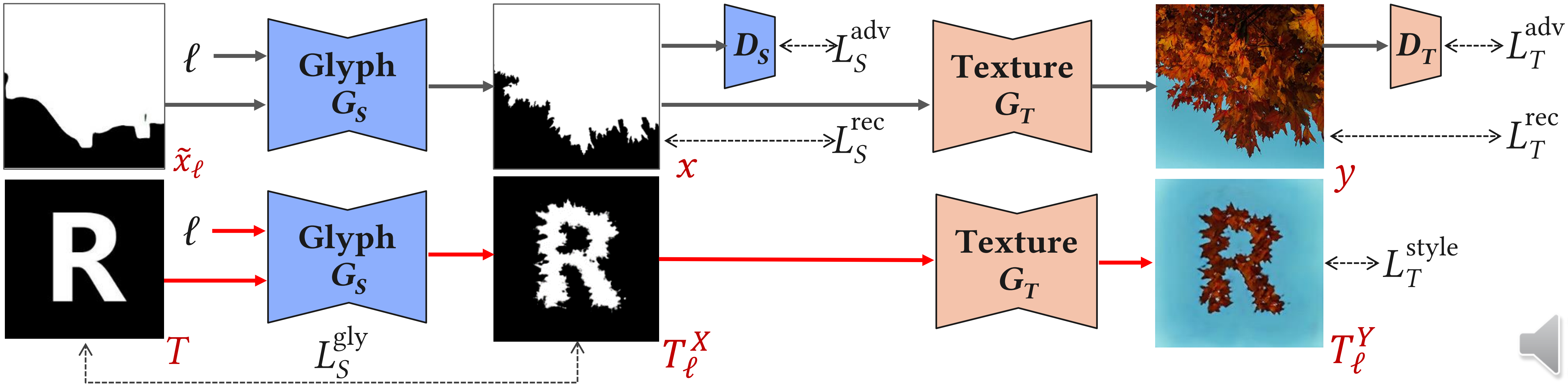
Forward Texture Transfer (G_T)

- Stage II: Forward Texture Transfer
  - Standard image-to-image translation framework
  - Train: learn to map x to y

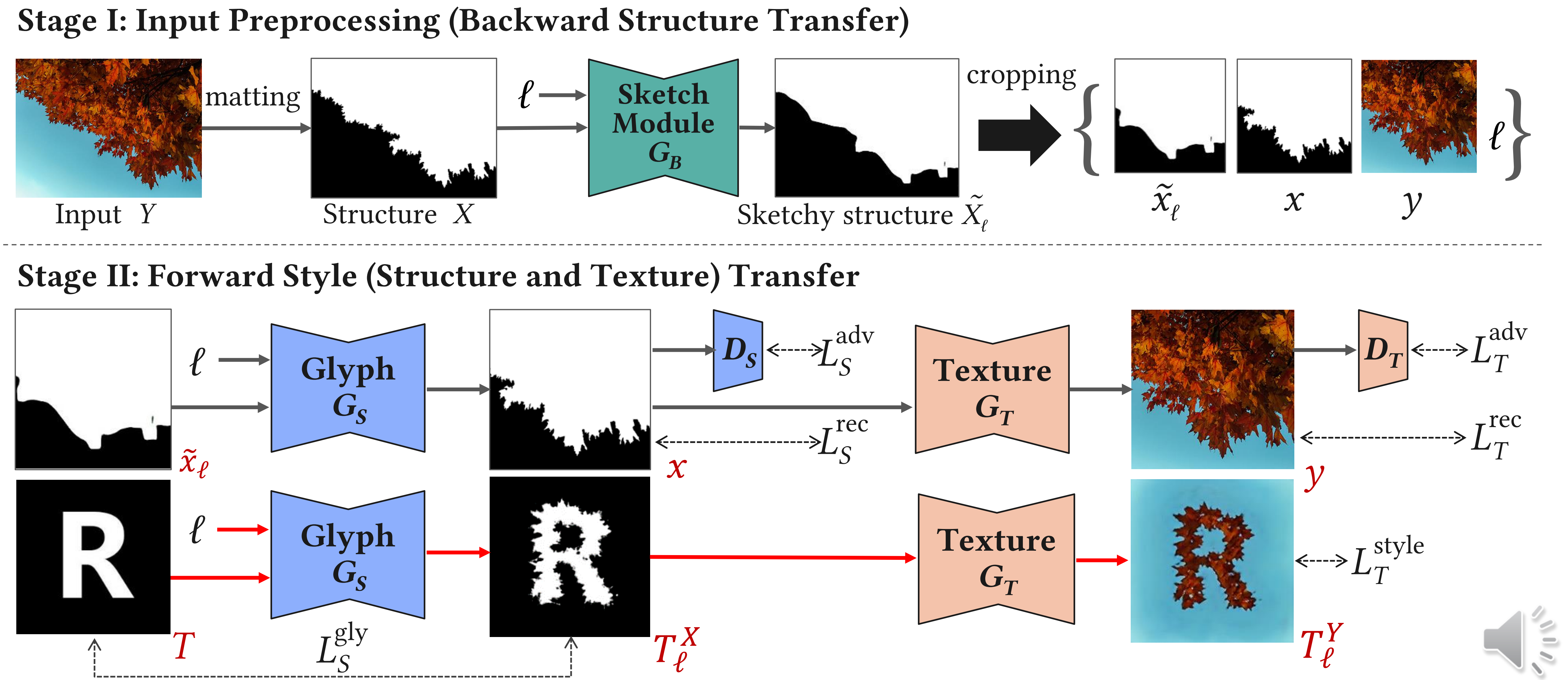


Forward Texture Transfer (G_T)

- Standard image2image translation framework
  - Test: render the texture in Y onto T_ℓ^X to yield the final artistic text T_ℓ^Y
  - Style loss: enhance texture details



Framework


Stage I: Input Preprocessing (Backward Structure Transfer)
matting
ℓ
Sketch Module G_B
cropping
Input Y
Structure X
Sketchy structure $\tilde{X}_\ell$
$\tilde{x}_\ell$
x
y
ℓ
Stage II: Forward Style (Structure and Texture) Transfer
ℓ
Glyph G_S
D_S
L_S^{adv}
L_S^{rec}
Texture G_T
D_T
L_T^{adv}
L_T^{rec}
$\tilde{x}_\ell$
x
y
ℓ
Glyph G_S
Texture G_T
L_T^{style}
T
L_S^{gly}
T_ℓ^X
T_ℓ^Y

Comparison with Other Methods

Input style | Target text | Image Analogy[1] | NST[2] | Doodle[3] | T-Effect[4] | UT-Effect[5] | Ours

[1] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin. Image analogies. SIGGRAPH. 2001

[2] L. A. Gatys, A. S. Ecker, and M. Bethge. Image style transfer using Convolutional neural networks. CVPR. 2016

[3] A. J. Champandard. Semantic style transfer and turning two-bit doodles into fine artworks. Arxiv. 2016

[4] S. Yang, J. Liu, Z. Lian, and Z. Guo. Awesome typography: statistics-based text effects transfer. CVPR. 2017

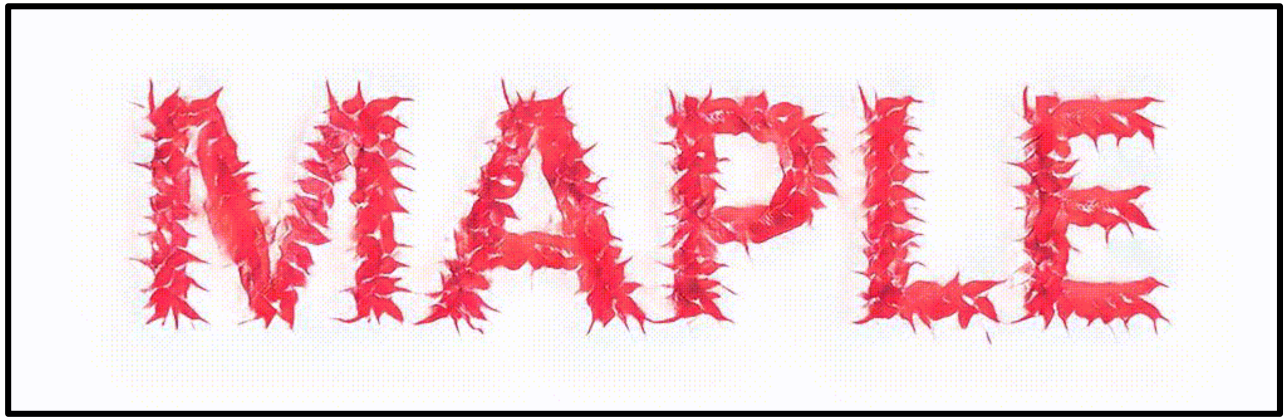
[5] S. Yang, J. Liu, W. Yang, and Z. Guo. Context-aware text-based binary image stylization and synthesis. TIP. 2019

Scale-Controllable Style Transfer

Reference style

MAPLE

Target text



legible → stylish

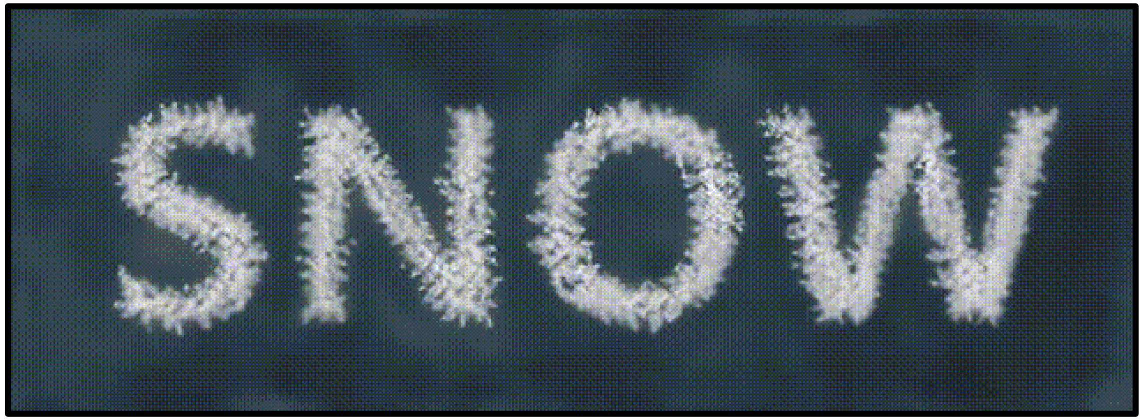
Adjusting glyph deformation degree

Scale-Controllable Style Transfer

Reference style

SNOW

Target text

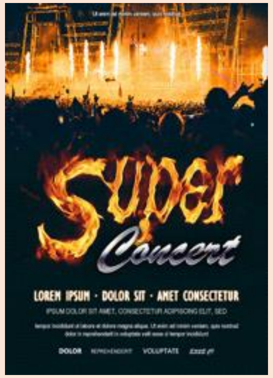


legible — stylish

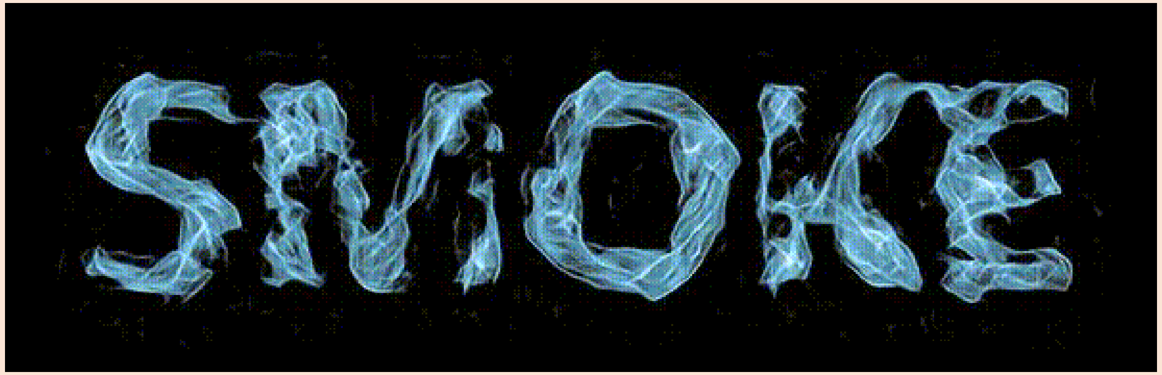
Adjusting glyph deformation degree

Applications

dynamic text generation



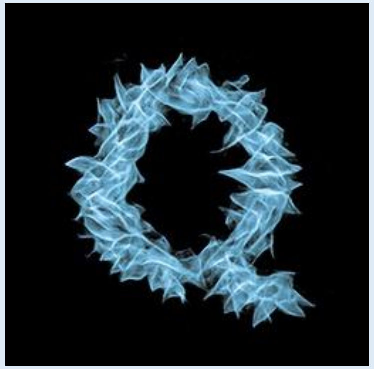
By adding random noises



By adding interpolated noise

diverse structure/texture mixture





stroke-based art design



- **Bidirectional Shape Matching**
  - Training data generation
    - Backward structure transfer
    - Image cropping
  - Fast forward multi-scale structure transfer
    - Smoothness-based sketch module
    - Controllable Resblock
- **Experimental Results**
  - Impressive results compared with other state-of-the-arts
  - Applications

Project

Poster Info:

Poster # 02

Poster 3.1 (Hall B)

Thursday, 10:30 – 13:00

Shuai Yang: http://williamyang1991.github.io