Mastering Sketching: Adversarial Augmentation for Structured Prediction

Edgar Simo-Serra*, Satoshi Iizuka*, Hiroshi Ishikawa (*equal contribution)

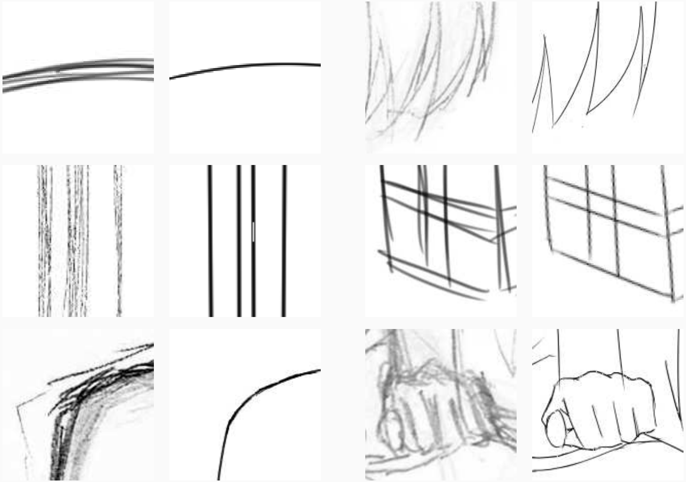
Wednesday, August 15, 2018

Waseda University


Menu
unnamed - 2016-02-24...
Thursday February 25, 00:32
unnamed - 2016-02-24_characters-of-pepper-carrot_007.kra [modified] - Krita - [unnamed - 2016-02-24_characters-of-pepper-carrot_007.kra]
File
Edit
View
Image
Layer
Select
Filter
Window
Tools
Settings
Help
Normal
Opacity: 1.00
Flow: 1.00
Size: 8.74 px
Advanced Color Selector
Brush Presets
All
Enter resource filters here
Layers
Normal
Opacity: 100%
line
fill
bg
sketch
background
Play Speed x300
Video: David Revoy, www.davidrevoy.com
deevad A5 pencil inking plus
RGB (8-bit integer/channel) sRGB built-in
5259 x 2400 (173M)
100%

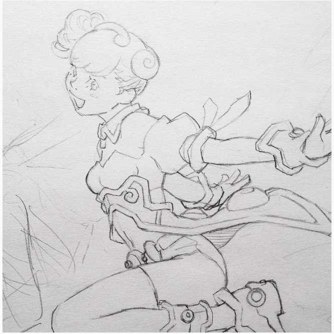
Sketch Simplification

Rough Target Rough Target



Contributions

- Semi-supervised framework for sketch simplification
- Pencil generation results
- Single-Image Optimization



Input

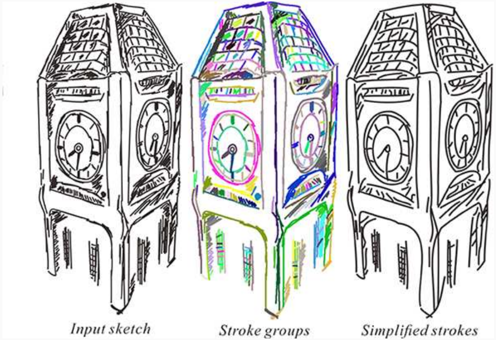
[Simo-Serra+ 2016]

Ours

Related Work

Related Work

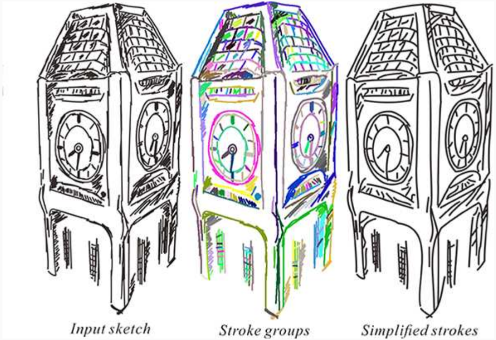
1. Sketch Simplification
   1.1 Progressive Online Modification
   1.2 Stroke Reduction
   1.3 Stroke Grouping



Liu et al. 2015

Related Work

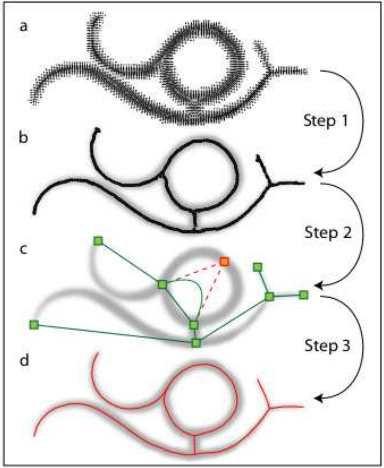
1. Sketch Simplification
    1.1 Progressive Online Modification
    1.2 Stroke Reduction
    1.3 Stroke Grouping
    1.4 Vector input



Liu et al. 2015

Related Work

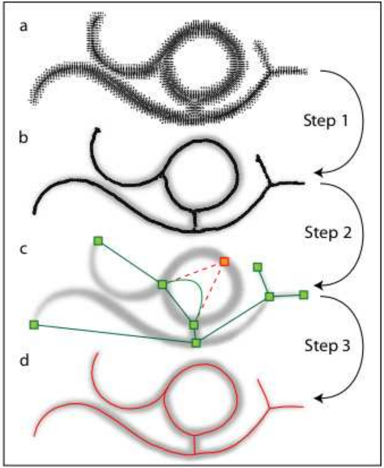
1. Sketch Simplification
   1.1 Progressive Online Modification
   1.2 Stroke Reduction
   1.3 Stroke Grouping
   1.4 Vector input
2. Vectorization
   2.1 Model Fitting (Bezier, ...)
   2.2 Gradient-based approaches



Noris et al. 2013

Related Work

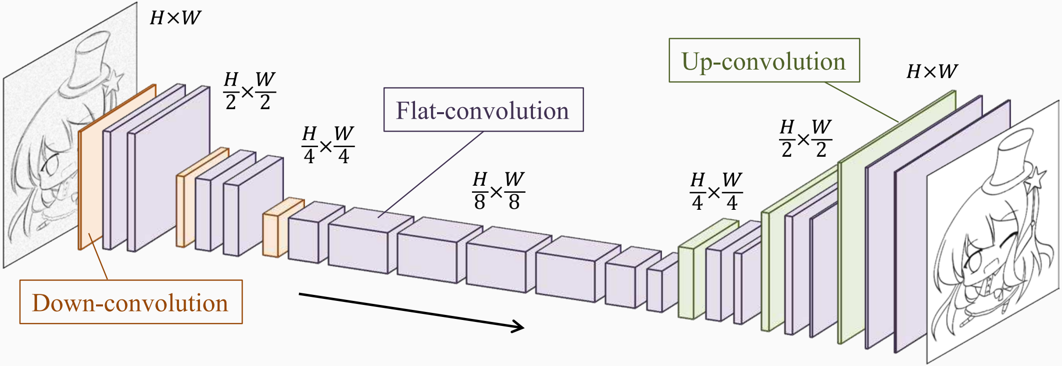
1. Sketch Simplification
    1.1 Progressive Online Modification
    1.2 Stroke Reduction
    1.3 Stroke Grouping
    1.4 Vector input
2. Vectorization
    2.1 Model Fitting (Bezier, ...)
    2.2 Gradient-based approaches
    2.3 Require fairly clean input sketches



Noris et al. 2013

Related work [Simo-Serra+ 2016]

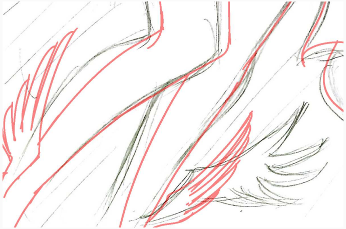
- 23 layer fully convolutional neural network
- Encoder-Decoder shape



Related work [Simo-Serra+ 2016]

- 23 layer fully convolutional neural network
- Encoder-Decoder shape
- Dataset construction is critical
- Expert knowledge is important

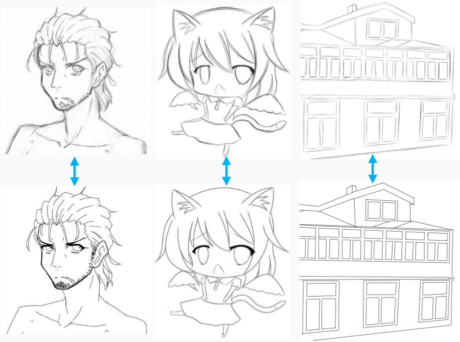
Standard Dataset Creation



Inverse Dataset Creation

Dataset Bias

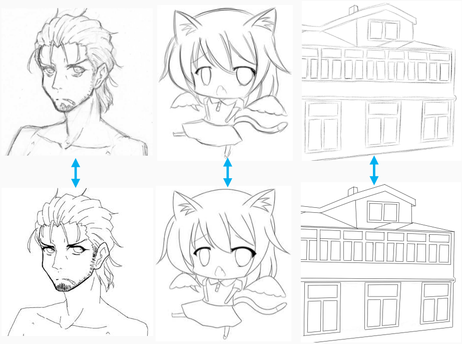
Training pairs



Rough sketches “in the wild”

Dataset Bias

Training pairs

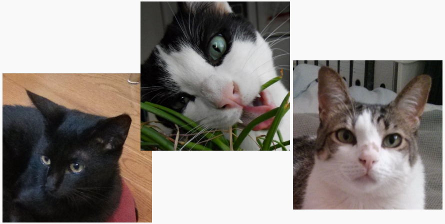


Rough sketches “in the wild”

- Supervised dataset (rough sketch and line drawing pairs): $\rho_{x,y}$
- Rough sketch dataset:ρ_x
- Line drawing dataset: ρ_y

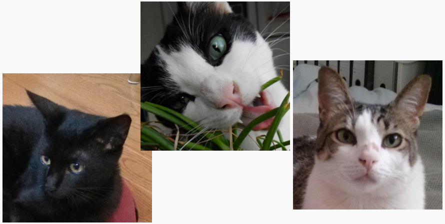
Generative Adversarial Networks

Generative Adversarial Network (GAN)



ρ_y

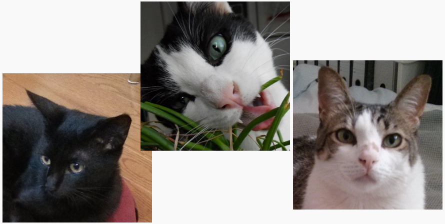
Generative Adversarial Network (GAN)



ρ_y

$$G \longrightarrow y$$

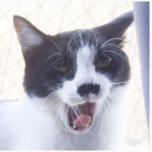
Generative Adversarial Network (GAN)



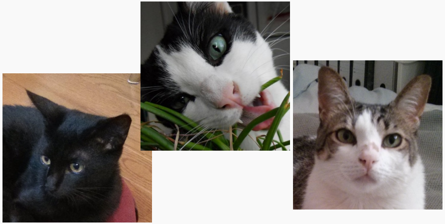
ρ_y

$$z \sim \mathcal{N}(0,1) \rightarrow \boxed{G} \rightarrow y$$

$$0.37 \rightarrow \boxed{G} \rightarrow$$



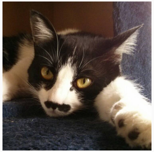
Generative Adversarial Network (GAN)



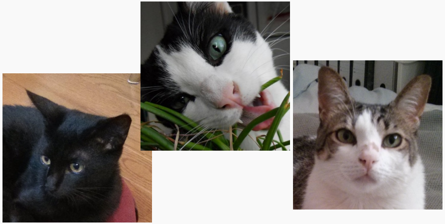
ρ_y

$$z \sim \mathcal{N}(0,1) \rightarrow \boxed{G} \rightarrow y$$

$$-0.01 \rightarrow \boxed{G} \rightarrow$$



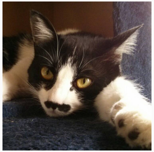
Generative Adversarial Network (GAN)



ρ_y

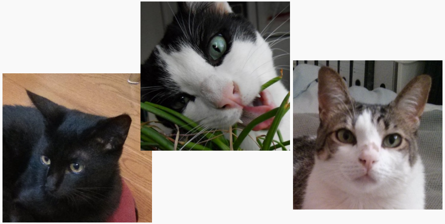
$$z \sim \mathcal{N}(0,1) \rightarrow \boxed{G} \rightarrow y \rightarrow \boxed{D} \rightarrow p(y \sim \rho_y)$$

$$-0.01 \rightarrow \boxed{G} \rightarrow$$



$$\rightarrow \boxed{D} \rightarrow 0.96$$

Generative Adversarial Network (GAN)



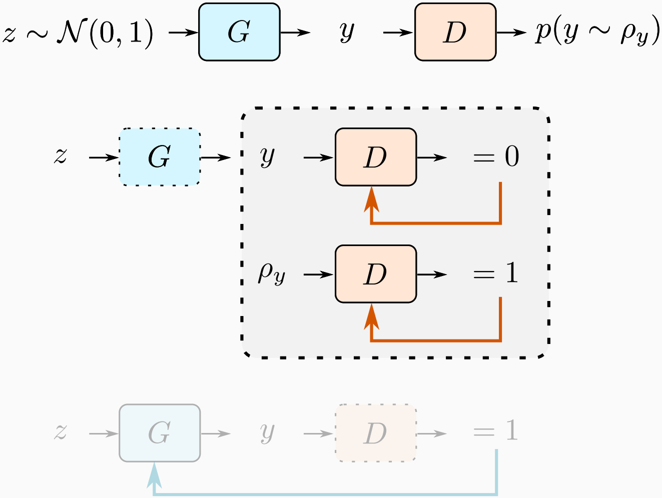
ρ_y

$$z \sim \mathcal{N}(0,1) \rightarrow \boxed{G} \rightarrow y \rightarrow \boxed{D} \rightarrow p(y \sim \rho_y)$$

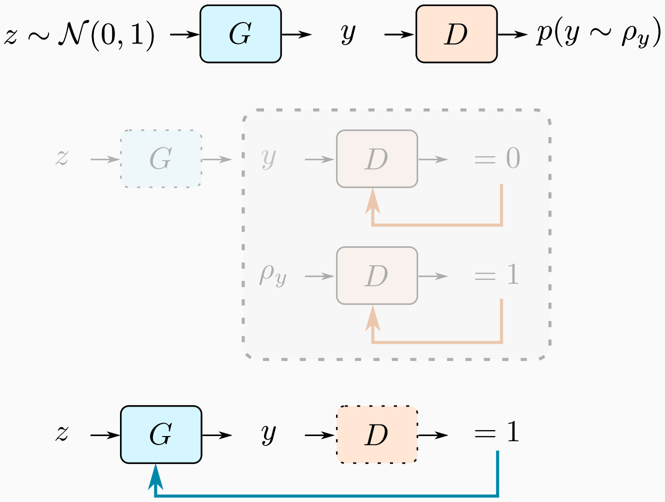
$0.07 \rightarrow \boxed{G} \rightarrow$

$\rightarrow \boxed{D} \rightarrow 0.01$

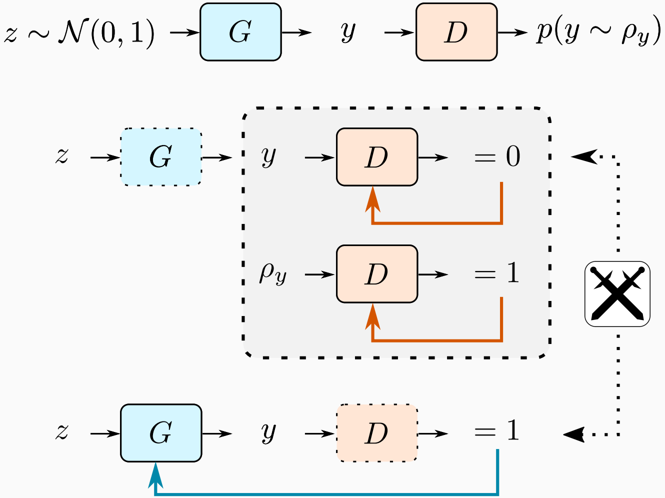
Generative Adversarial Network (GAN)



Generative Adversarial Network (GAN)



Generative Adversarial Network (GAN)



Generative Adversarial Network (GAN)

$$z \sim \mathcal{N}(0,1) \rightarrow \boxed{G} \rightarrow y \rightarrow \boxed{D} \rightarrow p(y \sim \rho_y)$$

- $D(\cdot)$: maximize classification prediction

$$\max_D \underbrace{\mathbb{E}_{y^*\sim\rho_y}}_{\text{Real data}} \log D(y^*) + \underbrace{\mathbb{E}_{z\sim\mathcal{N}(0,1)}}_{\text{Random}} \log(1 - D(G(z)))$$

Generative Adversarial Network (GAN)

$$z \sim \mathcal{N}(0,1) \rightarrow \boxed{G} \rightarrow y \rightarrow \boxed{D} \rightarrow p(y \sim \rho_y)$$

- $D(\cdot)$: maximize classification prediction
- $G(\cdot)$: minimize to fool $D(\cdot)$

$$\min_G \qquad \underbrace{\mathbb{E}_{z\sim\mathcal{N}(0,1)}}_{\text{Random}} \log(1 - D(G(z)))$$

Generative Adversarial Network (GAN)

$$z \sim \mathcal{N}(0,1) \rightarrow \boxed{G} \rightarrow y \rightarrow \boxed{D} \rightarrow p(y \sim \rho_y)$$

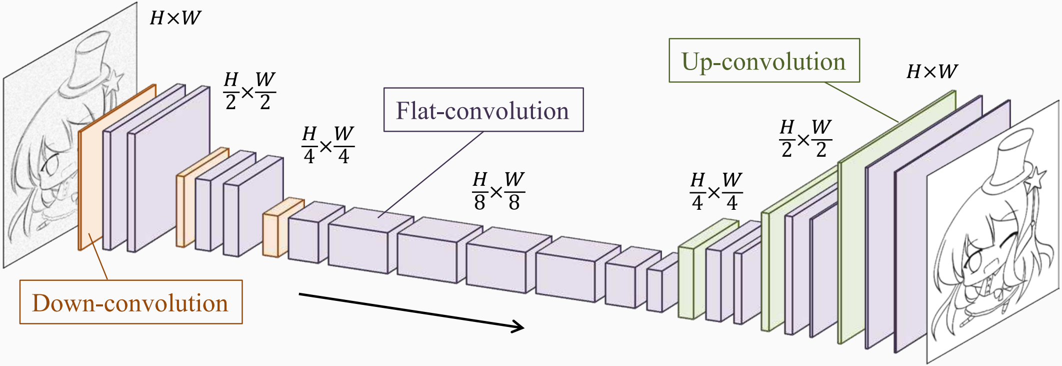
- $D(\cdot)$: maximize classification prediction
- $G(\cdot)$: minimize to fool $D(\cdot)$
- Alternate optimization

$$\min_G \max_D \underbrace{\mathbb{E}_{y^* \sim \rho_y}}_{\text{Real data}} \log D(y^*) + \underbrace{\mathbb{E}_{z \sim \mathcal{N}(0,1)}}_{\text{Random}} \log(1 - D(G(z)))$$

Framework

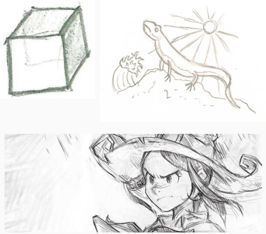
Model

- $S(\cdot)$: Sketch simplification model
  - 23 layer fully convolutional neural network [Simo-Serra+ 2016]
- $D(\cdot)$: Discriminator model
  - 6 layer convolutional neural network



Proposed framework

$$\min_{S} \overbrace{\mathbb{E}_{(x,y^*)\sim\rho_{x,y}}}^{\text{Supervised}} \overbrace{\|S(x) - y^*\|_2}^{\text{Standard Loss}}$$



入力

Standard Loss

Proposed framework

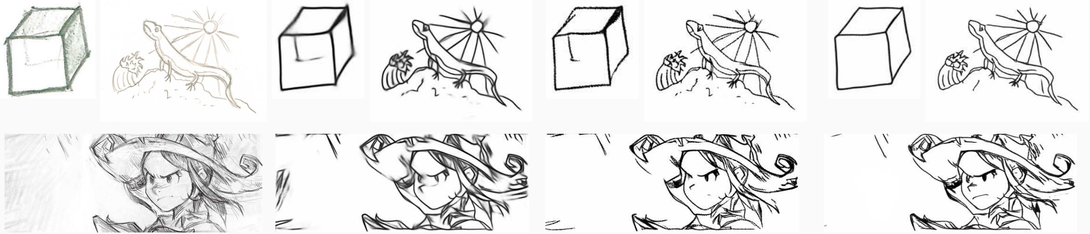
$$\min_S \max_D \overbrace{\mathbb{E}_{(x,y^*)\sim\rho_{x,y}}}^{\text{Supervised}} \left[\overbrace{\|S(x) - y^*\|_2}^{\text{Standard Loss}} + \overbrace{\alpha \log D(y^*) + \alpha \log(1 - D(S(x)))}^{\text{Adversarial Loss}} \right]$$

入力 Standard Loss +Adversarial

Proposed framework

$$\min_S \max_D \overbrace{\mathbb{E}_{(x,y^*)\sim\rho_{x,y}}}^{\text{Supervised}} \left[\overbrace{\|S(x)-y^*\|_2}^{\text{Standard Loss}} + \overbrace{\alpha\log D(y^*) + \alpha\log(1-D(S(x)))}^{\text{Adversarial Loss}} \right]$$

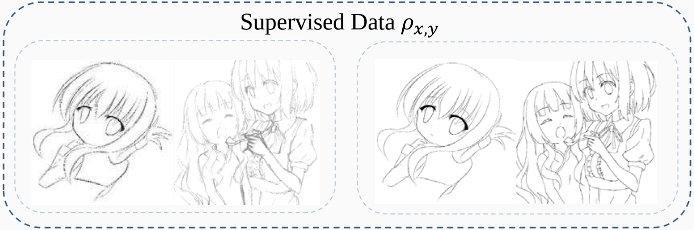
$$+ \underbrace{\beta\, \overbrace{\mathbb{E}_{y\sim\rho_y}}^{\text{Line}} [\, \log D(y)\,] + \beta\, \overbrace{\mathbb{E}_{x\sim\rho_x}}^{\text{Rough}} [\, \log(1-D(S(x)))\,]}_{\text{Unsupervised Adversarial Loss}}$$

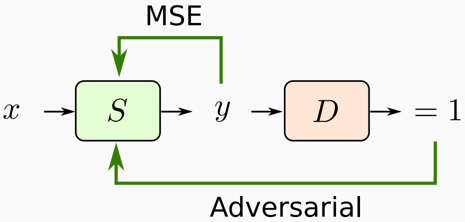


入力 | Standard Loss | +Adversarial | +Unsupervised

Training

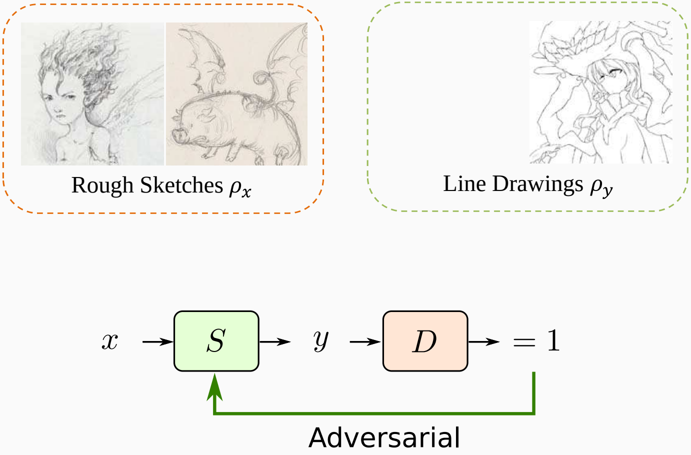
- Supervised data: standard loss + adversarial loss
- Unsupervised data: adversarial loss





Training

- Supervised data: standard loss + adversarial loss
- Unsupervised data: adversarial loss



Results

(lack of) Post-processing

- MSE loss requires post-processing to avoid blurring
- Adversarial loss avoids blurring

Input

LtS (no PP)

LtS (PP)

Ours (no PP)

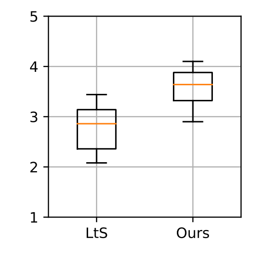
Results

Input [Favreau et al. 2016] [Simo-Serra et al. 2016] Ours

©Eisaku Kubonouchi

Perceptual User Study

- Comparison against LtS [Simo-Serra et al. 2016]
- 99 images, 15 users
- 94 images from artists not in training set
- 60 images come from twitter



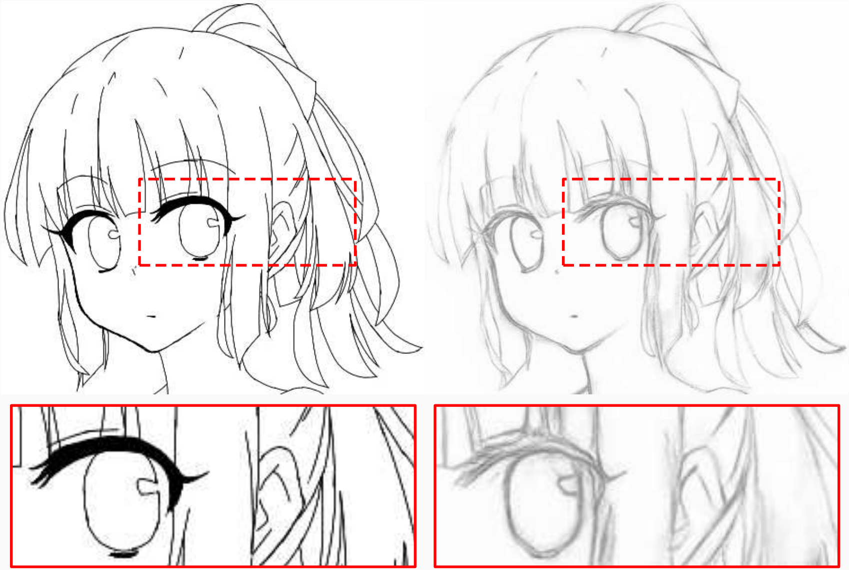
Absolute rating.

	LtS	Ours
absolute	2.77	3.60
vs LtS	-	88.9%
vs Ours	11.1%	-

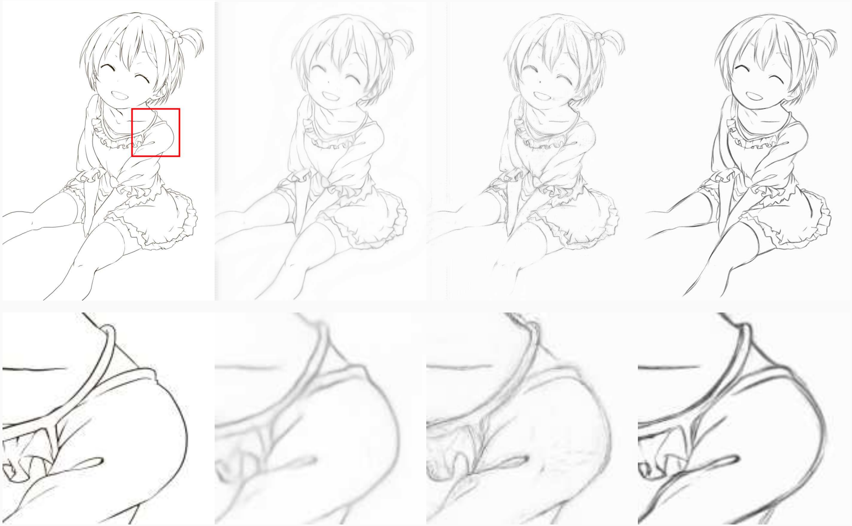
Mean results for all users.

Extensions

Extensions - Pencil Drawing Generation

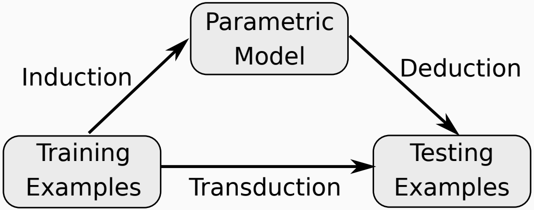


Extensions - Pencil Drawing Generation

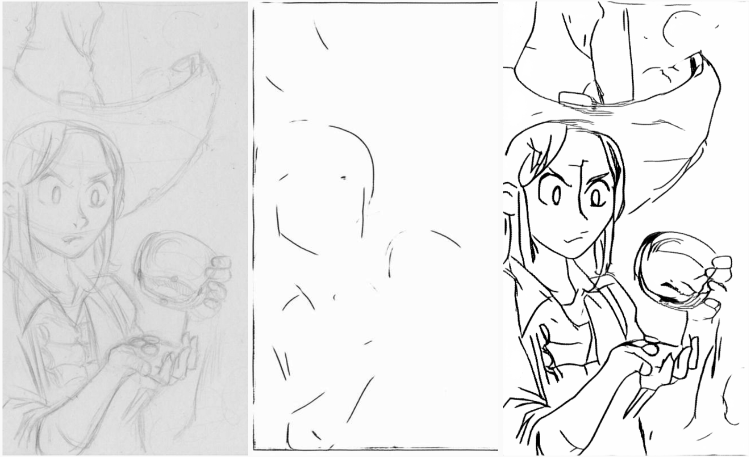


Input MSE Illustrator A Illustrator B

Extensions - Single-Image Optimization

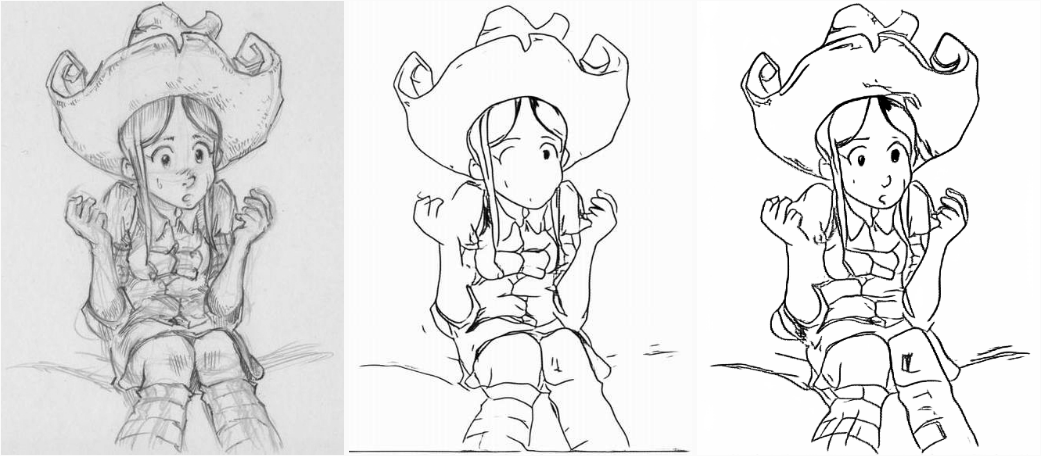


- **Inductive Machine Learning:** Learning a parametric function/model from training data and applying it on new data
- **Transductive Machine Learning:** Using training data to predict test data (model not necessary)



Input Ours Optimized

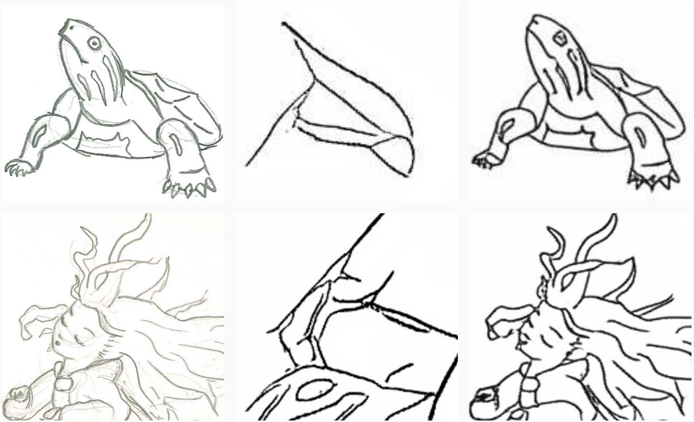
Extensions - Single-Image Optimization



Input Ours Optimized

Limitations

- Still a strong dependency on labelled data
- Results not perfect and require manual fixing

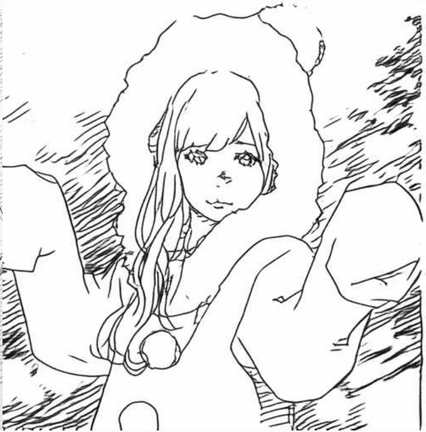


Input Only Unsupervised Ours

Limitations

- Still a strong dependency on labelled data
- Results not perfect and require manual fixing

Input



Output

To conclude

http://hi.cs.waseda.ac.jp/~esimo/research/sketch_master/

- Semi-supervised Sketch Simplification Framework
- Pencil Drawing Generation
- Single-Image Optimization

©Edgar Simo-Serra