Language-based Colorization of Scene Sketches

Changqing Zou*\textsuperscript{1,2}, Haoran Mo*\textsuperscript{1}, Chengying Gao\textsuperscript{1}, Ruofei Du\textsuperscript{3}, Hongbo Fu\textsuperscript{4}

Sun Yat-sen University\textsuperscript{1}
Huawei Noah’s Ark Lab\textsuperscript{2}
Google\textsuperscript{3}
City University of Hong Kong\textsuperscript{4}

Nov. 20\textsuperscript{th}, 2019
Motivation: Abstract Data and Human Cognition

- Sparse
- Highly abstract
Motivation: Abstract Data Understanding

- Lots of early exploration with computational models [Eitz et. al 2012, Li et. al 2013, Schneider et. al 2014, Li et. al 2015]
- Limited ability of understanding object-level sketches

Motivation: Sketch Understanding

TU-Berlin / Sketchy / QuickDraw
Motivation: Sketch Understanding

Scene-level sketch:
- Interaction among multiply objects
- More empty region, lack of contextual information

Motivation: Scene Sketch Understanding

Natural language

Scene sketch colorization
Motivation: Why Language-based?

- **Natural**: easily adopted by novice users
- **Touchless**: friendly for people with upper limb impairments
- **Effective**: support batch-processing colorization

“the bus is yellow with blue windows”
Motivation: Language-based Sketch Colorization

• Toy problem, but not simple ……
Challenges

A. Understanding scene-level sketch is very hard
   • Too abstract
   • Lack of contextual information

CMPlaces
[Castrejon et. al 2016]

Photo-Sketching
[Li et. al 2019]

SketchyScene
[Zou et. al 2018]
Challenges

B. Multimodal learning between language and scene sketch
   • Mapping between language and target objects

“the dog on the rightmost has orange body”
Challenges

B. Multimodal learning between language and scene sketch
   • Mapping between language and target objects
   • One or multiple objects with single instruction

“the two trees on the left of the house are light green”
Challenges

B. Multimodal learning between language and scene sketch

- Mapping between language and target objects
- One or multiple objects with single instruction
- Various free expressions of location

“the dog in the middle is gray”
/
“the dog near the house is gray”
Challenges

C. Multimodal learning between language and object sketch

- Object-part-level colorization
- Various free expressions of colors

“the bus is dark (navy/...) blue with white windows”
Inspiration: Drawing and Intelligence Development

- Sensitive to line drawing and color
- Mode of thinking and creation
Inspiration: Language and Literacy Development

- Embedding voice in traditional drawings supports children’s literacy development

[Raffle et. al 2007]
Related Work
A. Language-based Image Segmentation

- Fusion of textual and visual information
- Only natural images
- Only one binary mask for single or multiple target objects

[Ye et. al 2019]
Related Work

B. Language-based Image Colorization

- Language-based image editing (LBIE) [Chen et. al 2018]
- Require pair-wise scene-level sketch and color image

“The flower has red petals with yellow stigmas in the middle”

[Chen et. al 2018]

Our Work
Main contributions

- Language-based colorization **system** for scene sketches
- Language-based instance segmentation **network** for scene sketches
- Three large-scale **datasets** for language-based scene sketch colorization
Proposed approach

A. System pipeline
   - Divide-and-conquer and progressive strategy
   - Two modes (foreground and background)
   - Three models (instance matching, foreground colorization, background colorization)
Proposed approach

B.1 Instance Matching Model

- Training: two phases for binary mask (b) generation
- Inferring: fuse binary mask with instance segmentation results
Proposed approach

B.2 Foreground Colorization Model

- GAN + fusion module
- Colorize objects from different categories
Proposed approach

B.3 Background Colorization Model

- cGAN + two-branch decoder
  - Colorization branch
  - Explicit segmentation branch (segmentation loss)
Datasets

- **MATCHING dataset**: 38k groups of text-based instance segmentation data.

  - The bus has orange body and blue windows
  - The two trees on the left of the house are light green
  - All the clouds are dark gray
Datasets

- **MATCHING dataset**: 38k groups of text-based instance segmentation data.
- **_FOREGROUND dataset**: 4k groups of text-based sketch object colorization data.
Datasets

- **MATCHING dataset**: 38k groups of text-based instance segmentation data.
- **FOREGROUND dataset**: 4k groups of text-based sketch object colorization data.
- **BACKGROUND dataset**: 20k groups of text-based background colorization data.
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   • Single object

   “the bus is yellow with blue windows”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   • Single object

   “the bus is yellow with blue windows”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   • Multiple objects

   “all the trees are dark green”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions
     - Multiple objects
       “all the trees are dark green”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   • Colorize the background before all foregrounds

   “the sky is blue and the ground is green”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   • Colorize the background before all foregrounds

   “the sky is blue and the ground is green”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   • Language grammar error

   “the clouds are are in dark gray”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   "the clouds are are in dark gray"

   • Language grammar error
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   “the sun is yellow”
   “the bird on the left is red”
   “the bird on the right is dark brown”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

“the sun is yellow”
“the bird on the left is red”
“the bird on the right is dark brown”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   • Unsupported words

   “let the rabbit be in pink”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

• Unsupported words

“let the rabbit be in pink”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

“the duck on the right is orange”
“dark green grasses”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

“the duck on the right is orange”
“dark green grasses”
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   "colorize the bus in purple"
Results: Un-targeted

A. Un-targeted colorization experiment
   • Colorize a sketch with free instructions

   • Re-colorization

   “colorize the bus in purple”
Results: Un-targeted

A. Un-targeted colorization experiment

“the house is **orange** with **dark brown** roof”

“the two trees on the left are **dark green**”

“the three trees on the right are **dark green**”

“the sky is **cyan** and the ground is **gray**”

“let the dog to be **gray**”

“...”
Results: Un-targeted

A. Un-targeted colorization experiment

“the person on the left has red hair and is in dark brown shirt with light blue pants”

“the person on the right has red hair and is in orange shirt with yellow pants”

“colorize the sky purple and ground yellow”

“clouds are blue in the sky”
“grasses are green”

“the two ducks are yellow”
“the pig on the left is pink”
Results: Targeted

B. Targeted colorization experiment

- Colorize a sketch into target color images

```
"The sun is yellow"
"All the chickens are yellow"
"The house is red with dark brown roof and light blue windows"

"The sun is yellow with orange flame"
"All chickens are yellow with red crest and yellow feet"
"The walls of the house are brown and the roof of the house is red"
```
Results: Targeted

B. Targeted colorization experiment
   • Colorize a sketch into target color images

```
“the house is yellow with red roof”
“one duck on the left is purple”
“the other duck on the right is white”

“the house with red roofs has yellow doors”
“the left duck is purple”
“the right duck is white”
```
Results: Targeted

B. Targeted colorization experiment
   • Colorize a sketch into target color images

   “the leftmost bird is dark blue”
   “the bird on the right most is dark blue”
   “the two middle birds have blue body”
   “the birds are all blue”
Results: Generalization

C.1 Generalization experiment: cartoon-style drawings

“the person in the middle has dark brown hair and is in pink shirt with light gray pants”

“the person on the right has light brown hair and is in orange shirt with black pants”
Results: Generalization

C.1 Generalization experiment: cartoon-style drawings

“the rabbit on the right is **light brown**”

“the rabbit on the upper left is **dark gray**”
C.2 Generalization experiment: anime line art

“the person on the left has light brown hair and is in red shirt with dark gray pants”

“the person on the right has red hair and is in orange shirt with cyan skirt”

“the sky is pink and the ground is yellow”

“the person on the left has light brown hair and is in red shirt with dark gray pants”

“the person on the right has red hair and is in orange shirt with cyan skirt”

“the sky is pink and the ground is yellow”
Results: Generalization

C.3 Generalization experiment: artist freehand drawing

“the person on the left has red hair and is in yellow shirt with cyan pants”

“the person on the right has red hair and is in light brown shirt with purple pants”

“the sky is blue and the ground is green”
Results: Generalization

C.4 Generalization experiment: non-artist freehand sketches

"the dog is dark brown"

"the car is yellow with blue window"

"the house is blue with gray roof"

Sketchy
[Sangkloy et al. 2016]
Results: Generalization

C.4 Generalization experiment: non-artist freehand sketches

“all the trees are dark green”

“the dog on the right is dark brown”

“the sky is blue and the ground is green”

“the person is in light brown shirt with red pants”

Photo-Sketching [Li et al. 2019]
Limitations: Language Generality for Matching

“the bus is blue and the tree is light green”

“the tree is light green and the bus is blue”

• Multiple objects of different categories
Limitations: Language Generality for Matching

“the taxi is yellow with blue windows”

“the little boy has …”
“the little girl is in …”

• Alternative category names not in training data
Limitations: Language Generality for Colorization

- Arbitrary part-level information
- Arbitrary colors

“the wheels of the car is...”

“... blonde hair”
Limitations: Colorization Artifacts

Incorrect segmentation

Uncolored pixels
Limitations: Colorization Artifacts

Aliasing artifact
Future work: Multimodal Colorization System

- Language-based: more natural and accessible
- Scribble-based: direct and precise control
To conclude

• Human’s understanding of abstract data at scene level.

• The first language-based colorization system for scene sketches.

• Three large-scale datasets for language-based scene sketch colorization.

• Plausible results with room for improvement.
Dataset and code

- Project Page: https://sketchyscene.github.io/SketchySceneColorization/
- Code: https://github.com/SketchyScene/SketchySceneColorization
- Lab. Homepage: http://sysu-imsl.com/

Acknowledgments

- Participants on data annotations
- Reviewers
- Guangzhou Science Technology and Innovation Commission
- City University of Hong Kong

Thank you!