Sketch to Photo Translation

Yuhang Li, Xuejin Chen, Xiangxiang Wang, Sing Bing Kang

Original idea

Sketch to Photo Translation

- Problem definition
- Why sketch?
- Characteristic of Sketch

Related works

- Sketch2Photo
- Sketch-based classification
- Sketch-based retrieval
- Image-to-image translation

Exploration

- CycleGAN
- DiscoGAN

Original idea

Sketch to Photo Translation

- Problem definition
- Why sketch?
- Characteristic of Sketch

Related works

- Sketch2Photo
- Sketch-based classification
- Sketch-based retrieval
- Image-to-image translation

Exploration

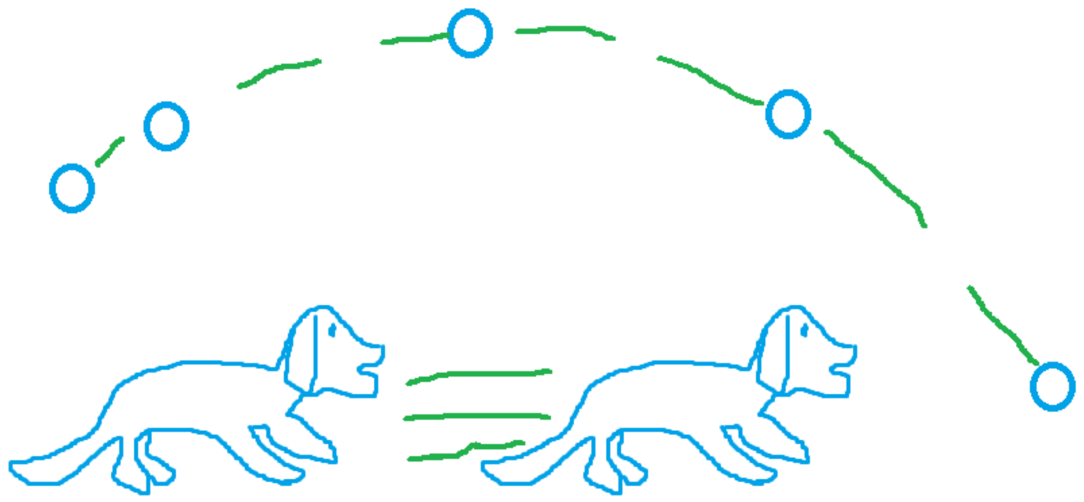
- CycleGAN
- DiscoGAN

Original idea

Sketch to realistic video

Two kinds of strokes in a sketch:

- Object strokes
  - What object in the scene
- Motion strokes
  - How the object moves in the scene

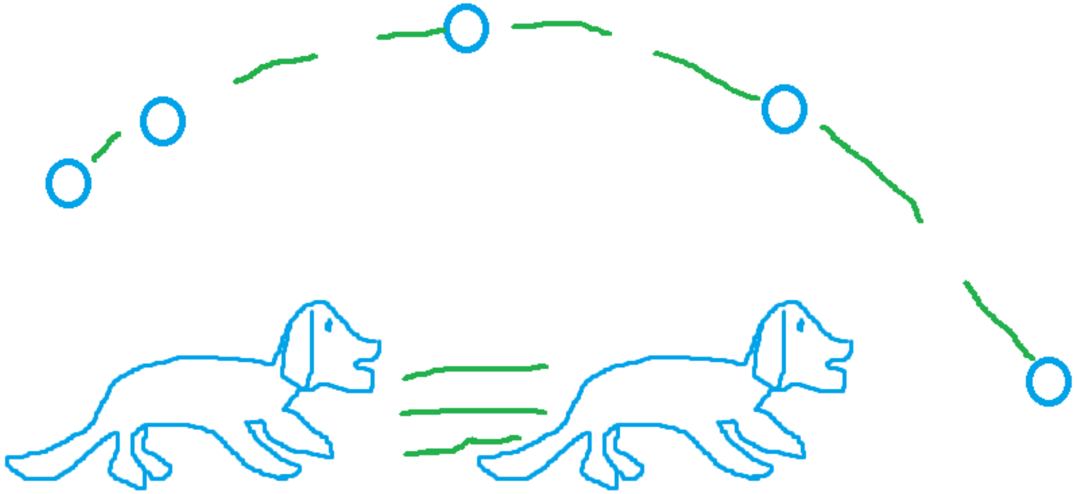


Original idea

Sketch to realistic video

Two kinds of strokes in a sketch:

- Object strokes
  - What object in the scene
- Motion strokes
  - How the object moves in the scene



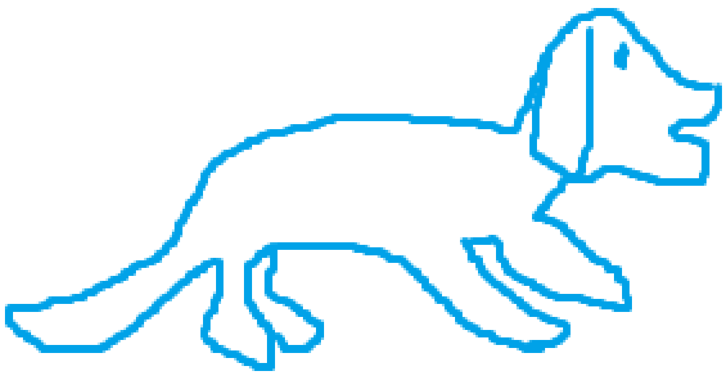
It is difficult to obtain corresponding videos that match both object strokes and motion strokes for training!

Original idea

Sketch to realistic video

Two kinds of strokes in a sketch:

- Object strokes
  - What object in the scene
- Motion strokes
  - How the object moves in the scene



It is difficult to obtain corresponding videos that match both object strokes and motion strokes for training!

Focus on object strokes!

Original idea

Sketch to Photo Translation

- Problem definition
- Why sketch?
- Characteristic of Sketch

Related works

- Sketch2Photo
- Sketch-based classification
- Sketch-based retrieval
- Image-to-image translation

Exploration

- CycleGAN
- DiscoGAN

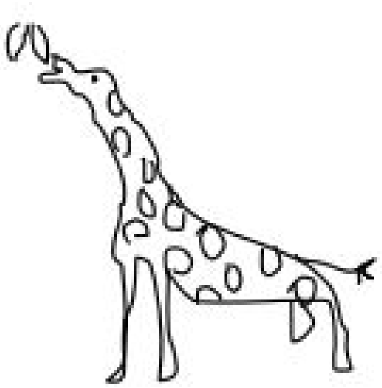
Problem definition

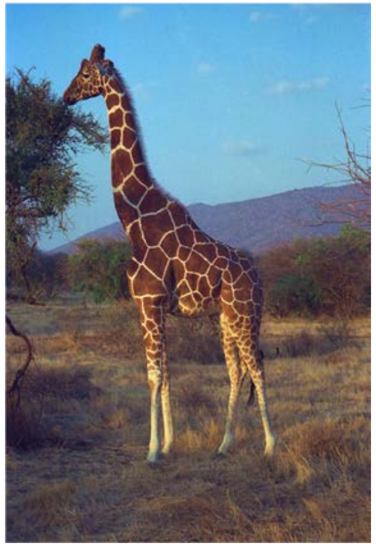
Sketch-based image generation

Given a sketch of **one** object, we want to generate a **new** but **realistic** image that preserves (Discuss)

- Object class: giraffe, elephant, chair or car
- View point
- Layout: size of the object, position in the image
- Pose: standing or sitting
- Background ?

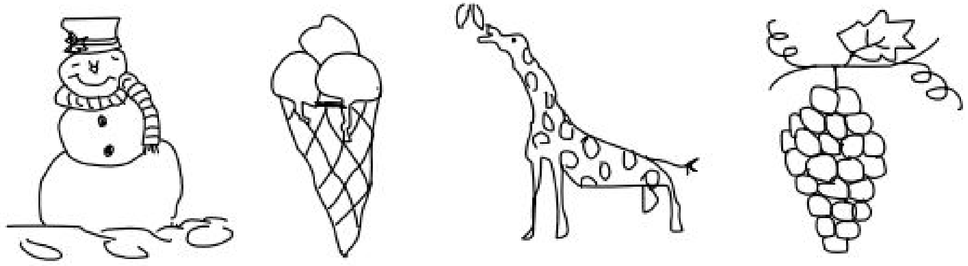
Shared information!





Sketch: why sketch?

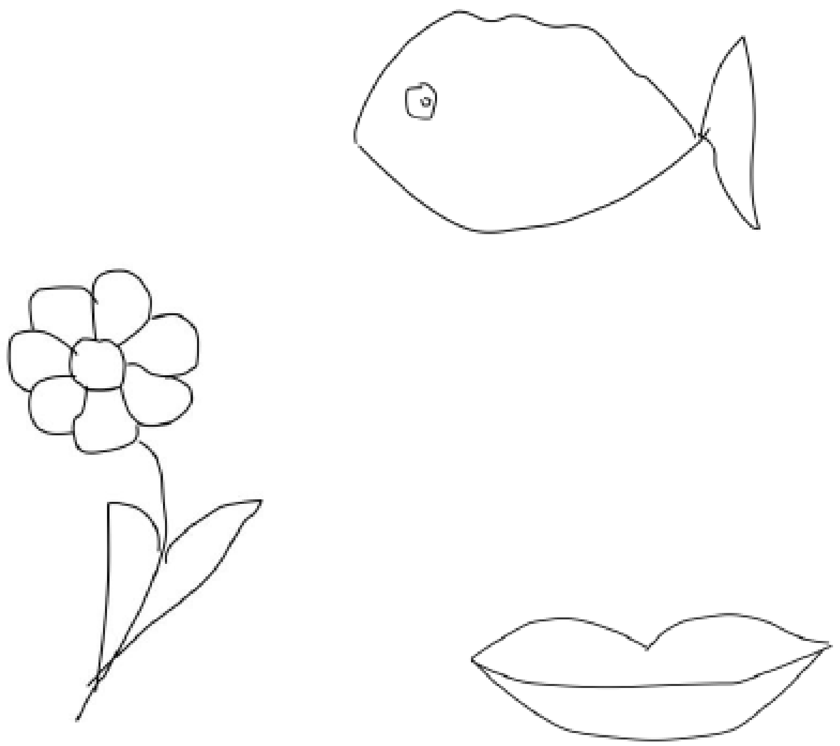
- A universal form of communication across nations and cultures
- Tracing back to prehistoric cave painting
- Conveying abstract concepts visually
- A *sketch* can speak a thousand words
- Becoming more important due to the popularity of touch devices



Characteristic of Sketch

- Iconic

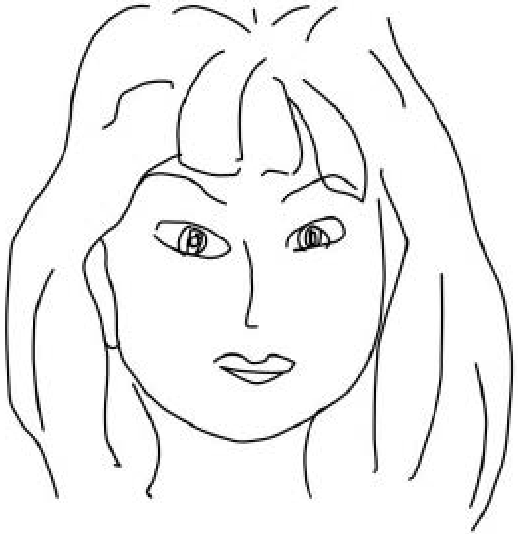
  Sketches can be highly abstract.



Characteristic of Sketch

- High intra-class variation

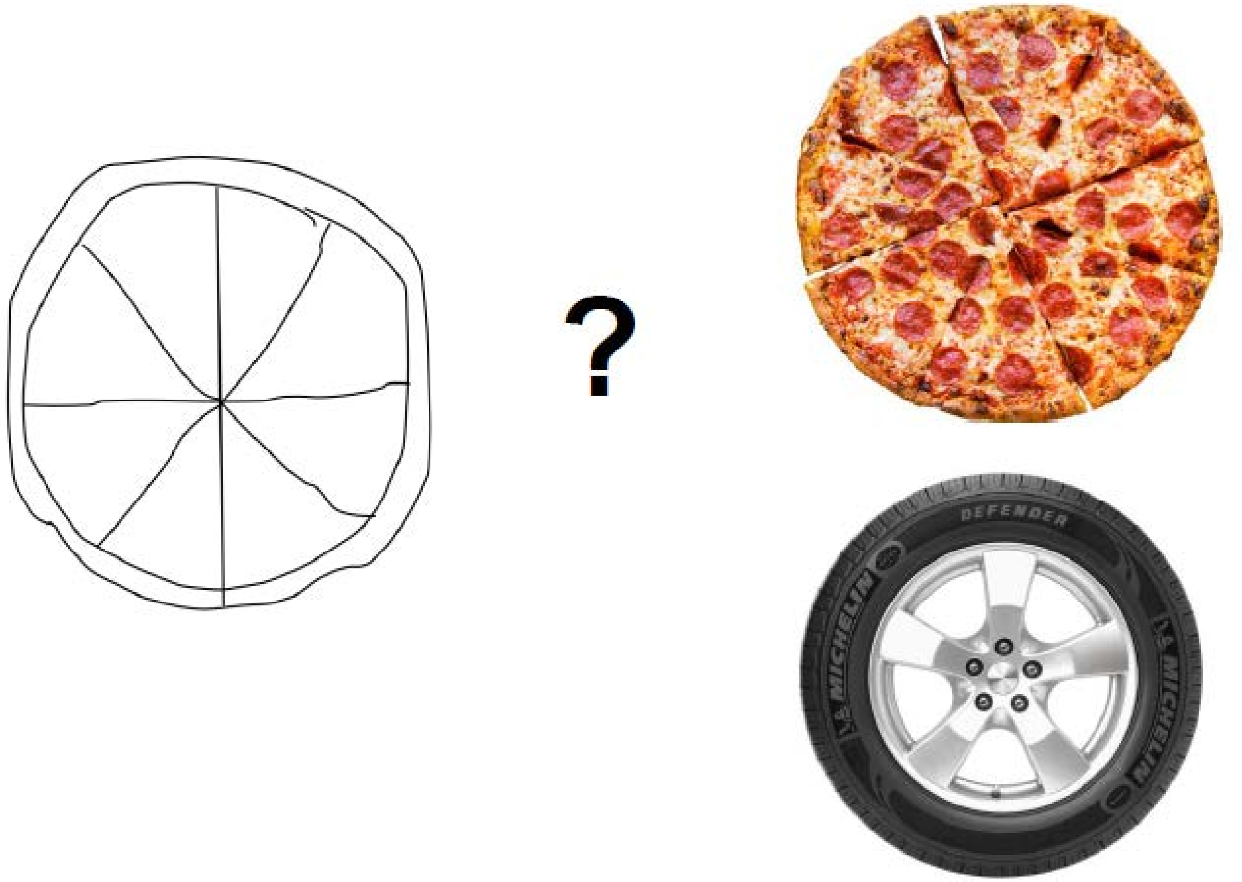
  Sketches drawn by different people vary in the level of abstraction or deformation.



Characteristic of Sketch

- Lack of visual cues

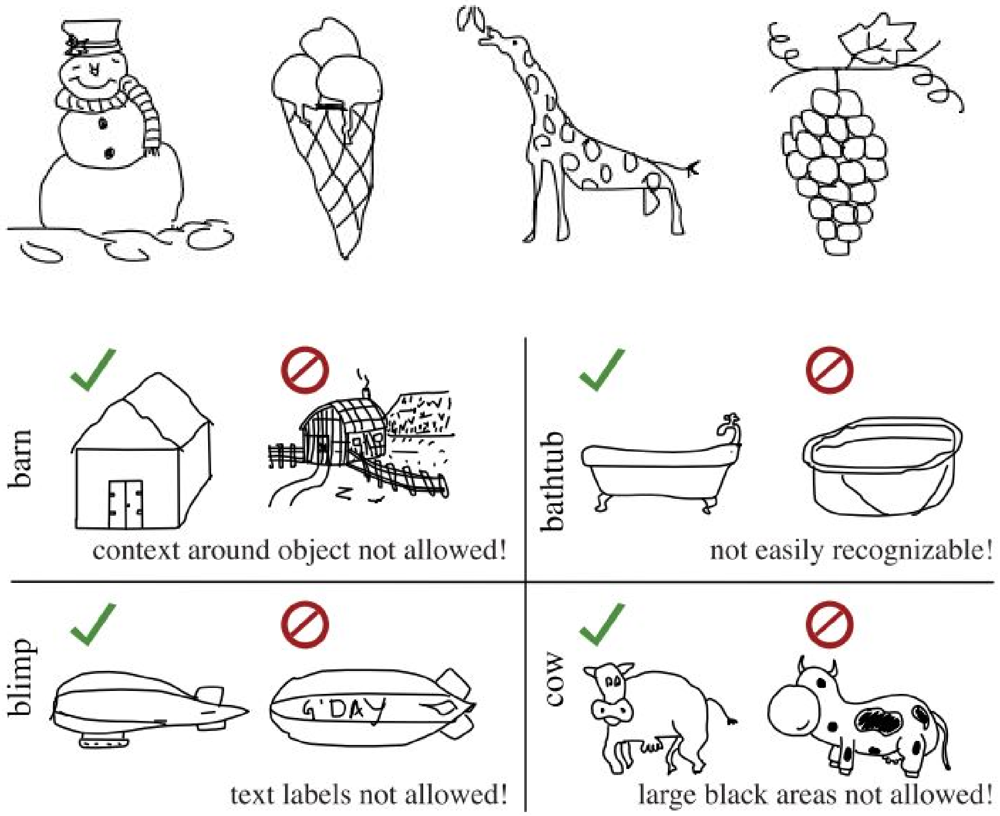
  Sketches are lack of colors and textures compared to photos.



Dataset

- TU-Berlin sketch dataset: a sketch dataset containing 250 categories and 80 sketches in each category.
- Fine-grain sketch dataset: paired fine-grain sketch and image of shoes and chairs.

Fine-grain sketch dataset



TU-Berlin sketch dataset

Original idea

Sketch to Photo Translation

- Problem definition
- Why sketch?
- Characteristic of Sketch

Related works

- Sketch2Photo
- Sketch-based classification
- Sketch-based retrieval
- Image-to-image translation

Exploration

- CycleGAN
- DiscoGAN

Sketch2Photo: Internet Image Montage

- Generate a realistic image using a simple freehand sketch annotated with text labels.
- Text labeled sketch
- Seamlessly stitching several photograph

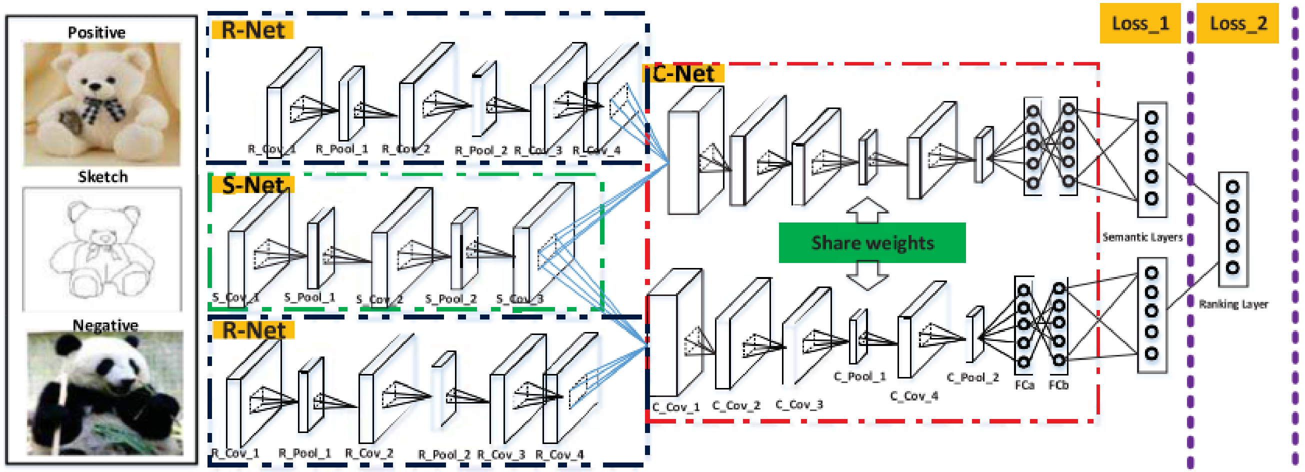
Sketch2Photo: Internet Image Montage

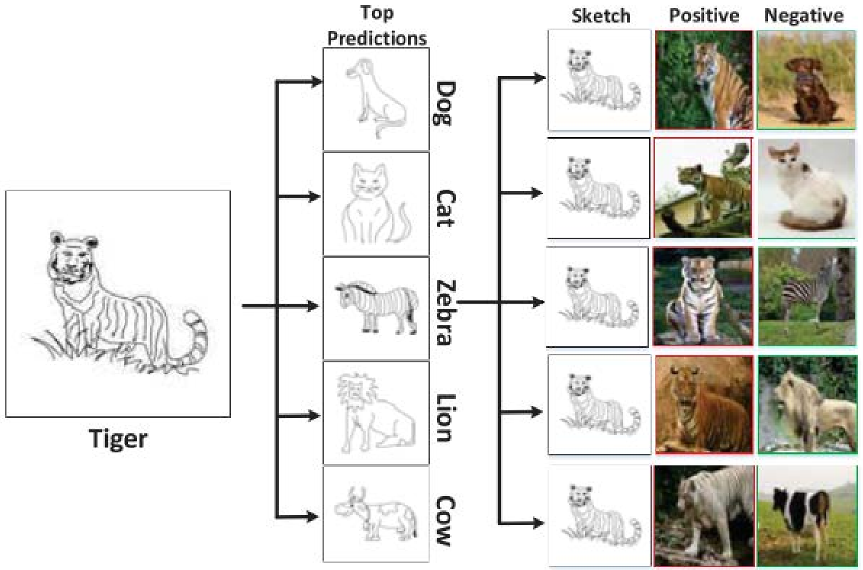
- Generate a realistic image using a simple freehand sketch annotated with text labels.

- Text labeled sketch
  - Labeling should be done automatically! (sketch classification)
- Seamlessly stitching several photograph
  - We want generate a totally **new** image, rather than a combination of existing image parts.

Sketch classification: SketchNet

- Discover the discriminative structures of sketch images
- Triplet networks based on Siamese nets
- Weekly supervised: with the help of photo images

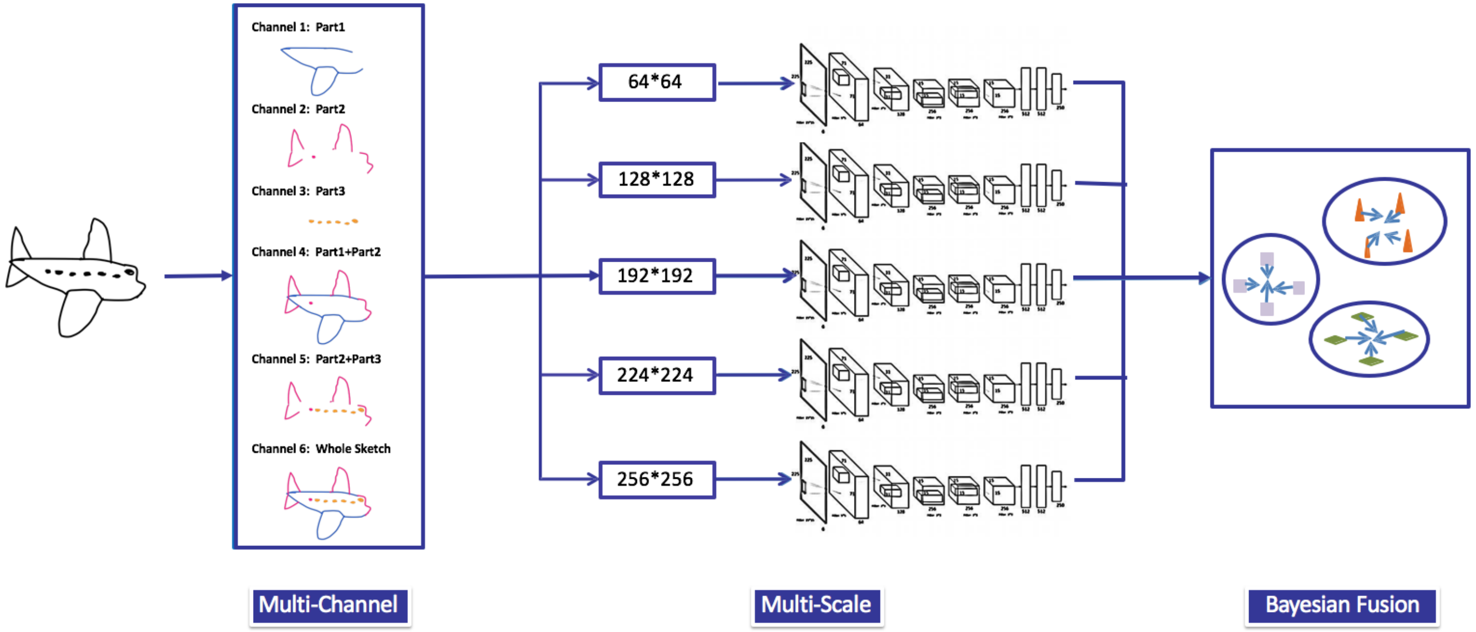




H. Zhang *2016*

Sketch classification: Sketch-a-Net

- A novel method of data augmentation
- Bayesian fusion
- Multiple models of different scales
- Beats human in sketch classification performance



Sketch-a-Net, Q. Yu 2016

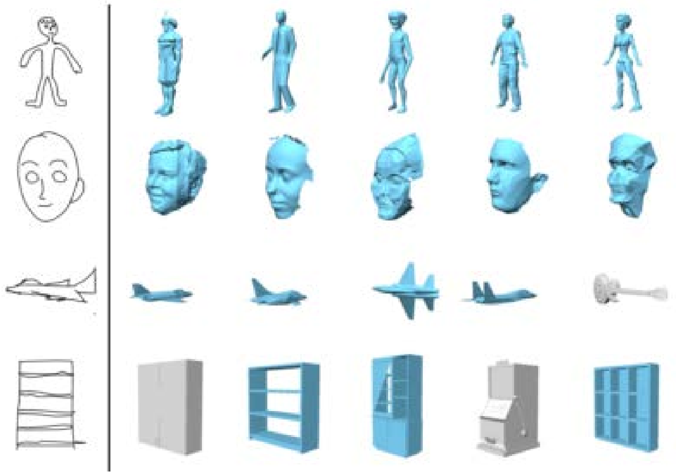
Sketch classification

Summary and inspirations

- Deep neural networks are able to handle sketch classification.
- Sketches with high intra-class variance are still able to mapped together in specific embedding space.
- There should be no need to annotate the sketch with text label.

Sketch-based retrieval

- Sketch-based 3D model retrieval
- Sketch-based image retrieval
- Sketch-based fine-grain image retrieval



3D model retrieval, F. Wang et al 2015

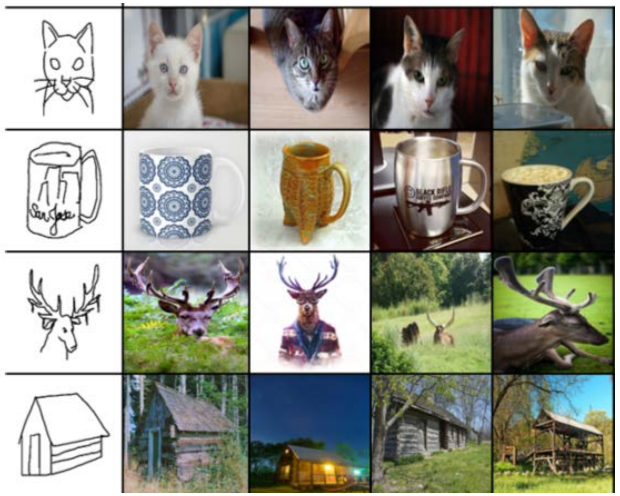


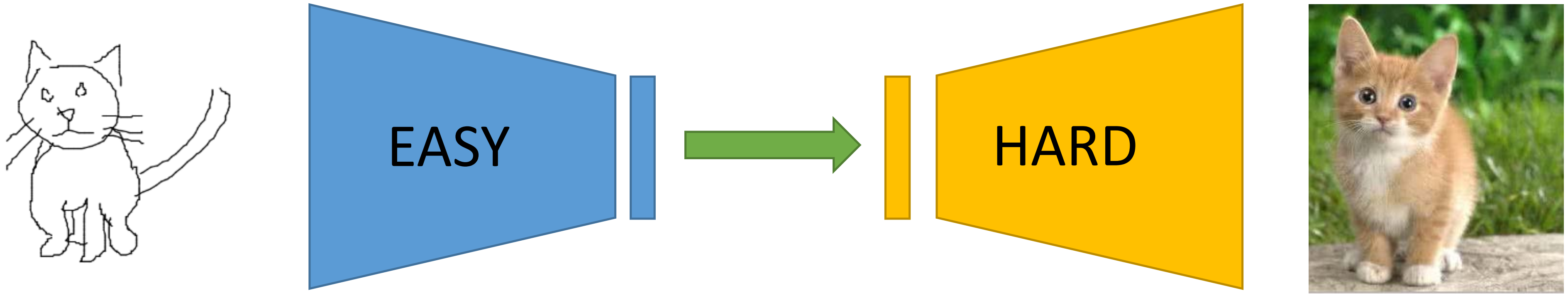
Image retrieval, P. Sangkloy et al 2016

Fine-grain image retrieval, Q. Yu et al 2016

Sketch-based retrieval

Summary and inspirations

- Minimize the distance between sketch and photo in embedding space
- A sketch and a similar photo are able to mapped close to each other in specific embedding space.



Inspired architecture of sketch-to-photo translation

Image-to-image translation

- Based on generative adversarial nets (GANs)
- Paired images for training
- A framework for multiple applications, only switching the training sets.

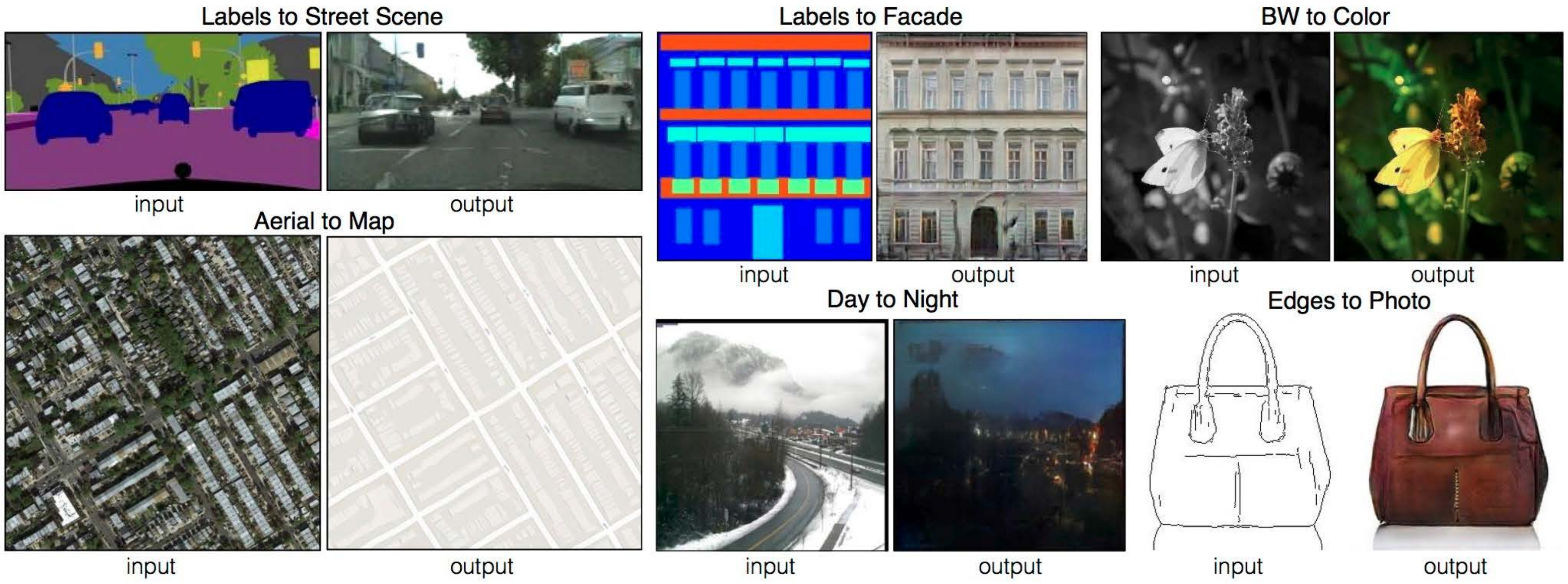


Image-to-image translation, P. Isola et al 2016

Image-to-image translation

- Based on generative adversarial nets (GANs)
- Paired images for training
- A framework for multiple applications, only switching the training sets.

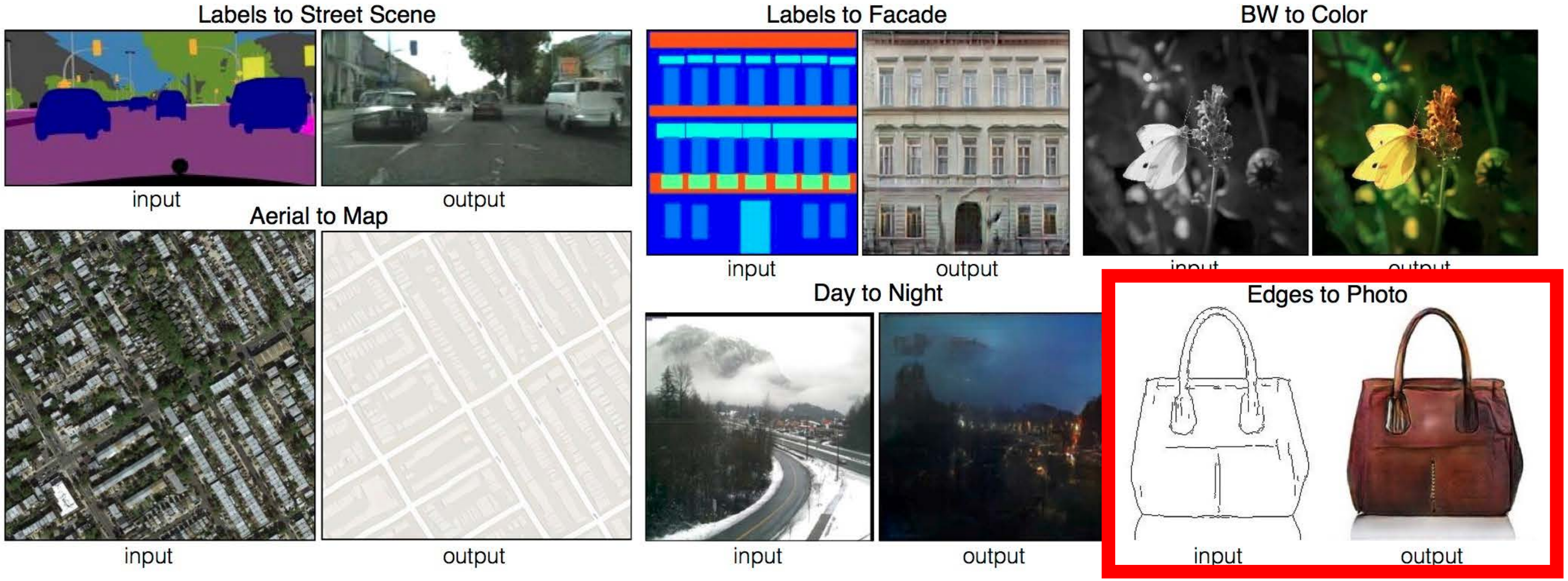


Image-to-image translation, P. Isola et al 2016

Image-to-image translation

Edge map to photo, P. Isola et al 2016

Image-to-image translation

- Paired images are required
  - Highly complex, expensive to obtain
  - Sometimes, not even well–defined
- Generated images are aligned to the edges. (Not only in the application of edge-to-photo)
- Not able to generate good results from abstract sketches

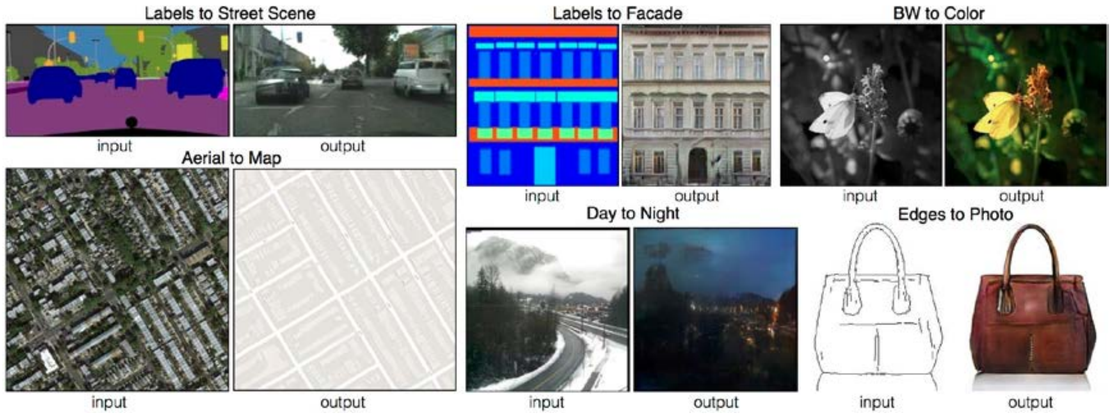


Image-to-image translation, P. Isola et al 2016

Input abstract sketch   Output image   Expected output

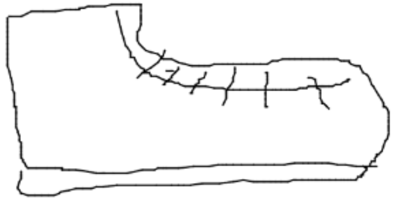
Image-to-image translation

- Paired images are required
  - Highly complex, expensive to obtain
  - Sometimes, not even well–defined
- Generated images are aligned to the edges. (Not only in the application of edge-to-photo)
- Not able to generate good results from abstract sketches

Edge map to photo, P. Isola et al 2016

Input edge map

Good result



Input abstract sketch

Bad result

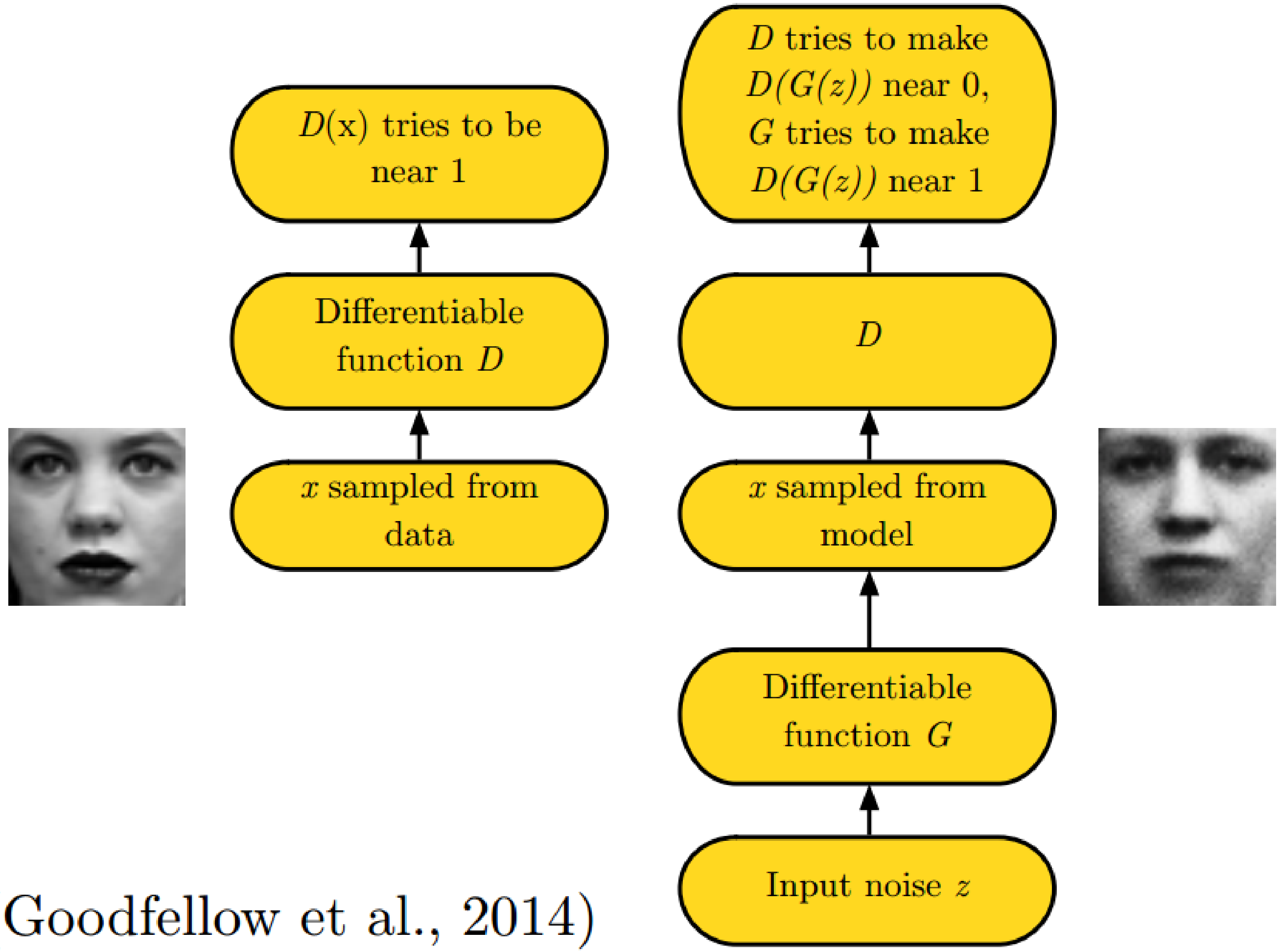
Image-to-image translation

Summary and inspirations

- GAN is a powerful framework to generate images.
- Images generated by image-conditional GANs tend to remember the edge maps of the input images.

Generative Adversarial Nets (GANs)

- Generate images of a given dataset by adversarial training
- Discriminator
- Generator
- Minmax game

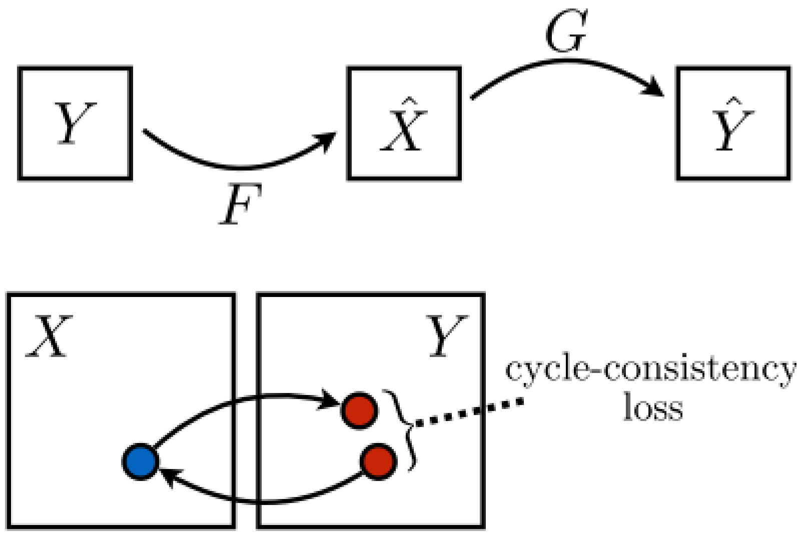


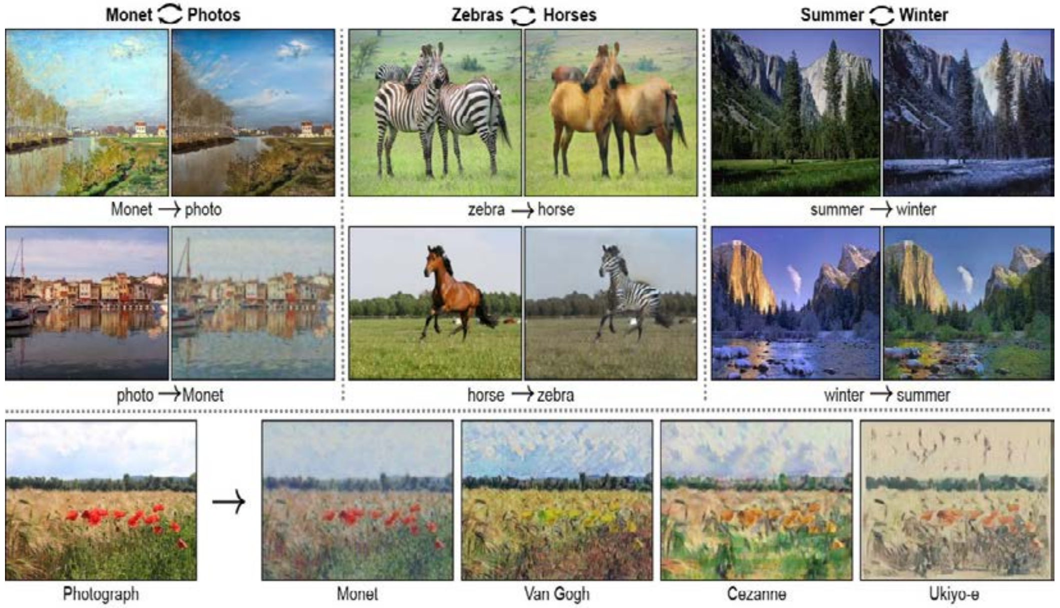
(Goodfellow et al., 2014)

Unpaired image translation

- Unpaired images datasets
- Cycle consistency

  Reconstruct the input image from generated image



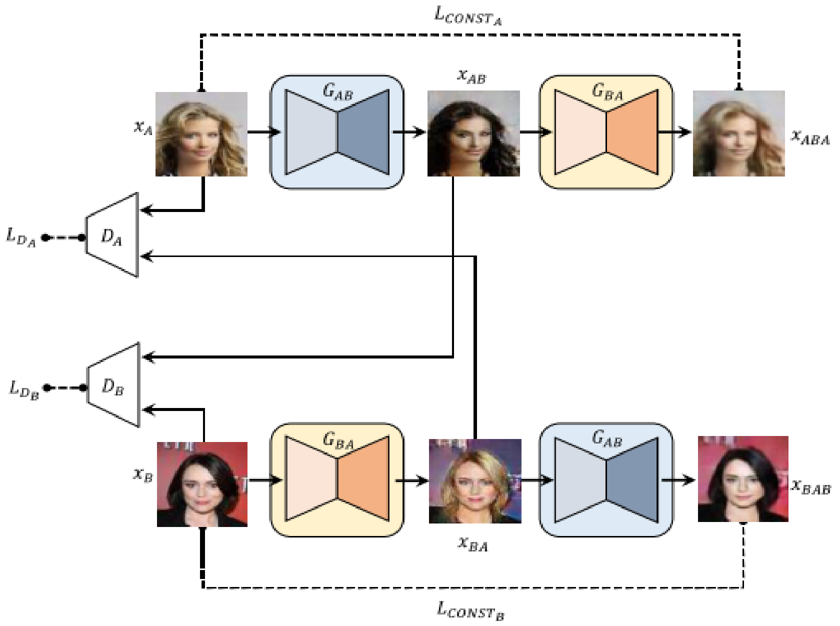


CycleGAN, J. Zhu et al 2017

Unpaired image translation

- Unpaired images datasets
- Cycle consistency (same idea)

  Reconstruct the input image from generated image



(a) Learning cross-domain relations **without any extra label**

(b) Handbag images (input) & **Generated** shoe images (output)

(c) Shoe images (input) & **Generated** handbag images (output)

DiscoGAN, T. Kim et al 2017

Original idea

Sketch to Photo Translation

- Problem definition
- Why sketch?
- Characteristic of Sketch

Related works

- Sketch2Photo
- Sketch-based classification
- Sketch-based retrieval
- Image-to-image translation

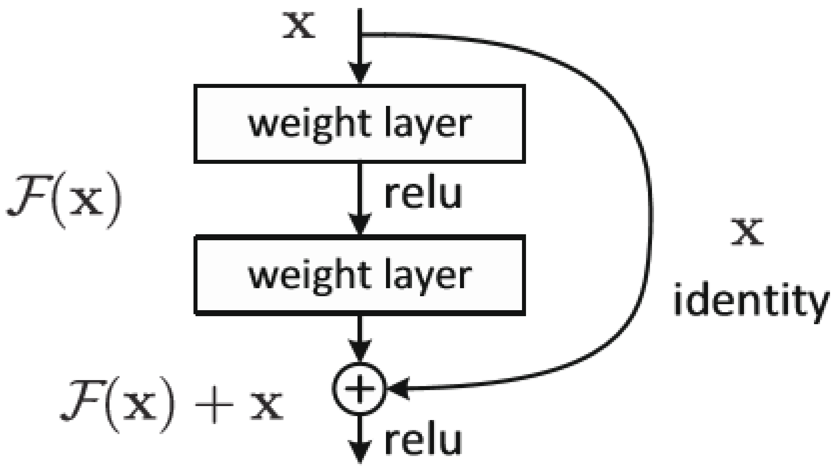
Exploration

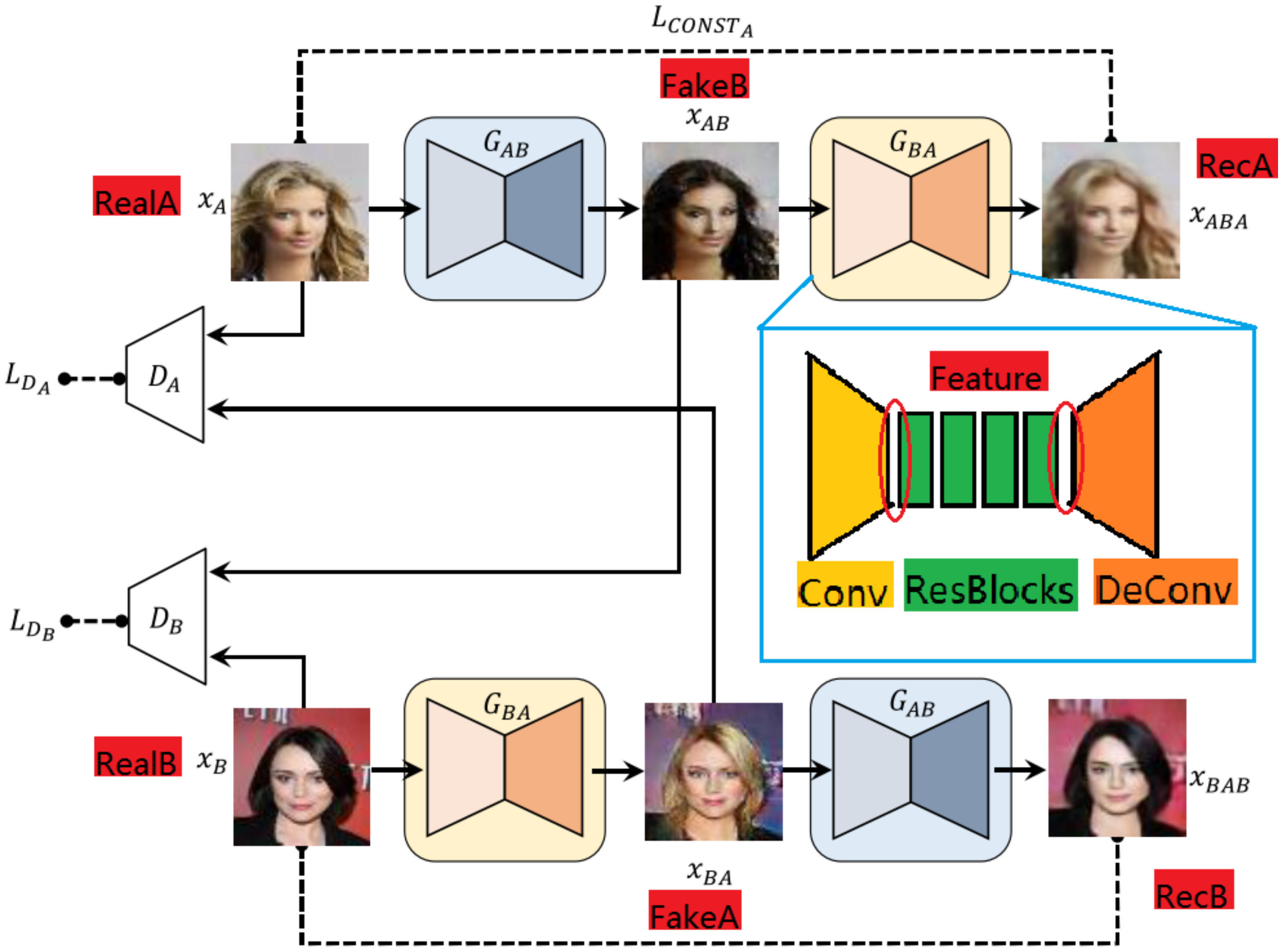
- CycleGAN
- DiscoGAN

Exploration: CycleGAN

Architecture details

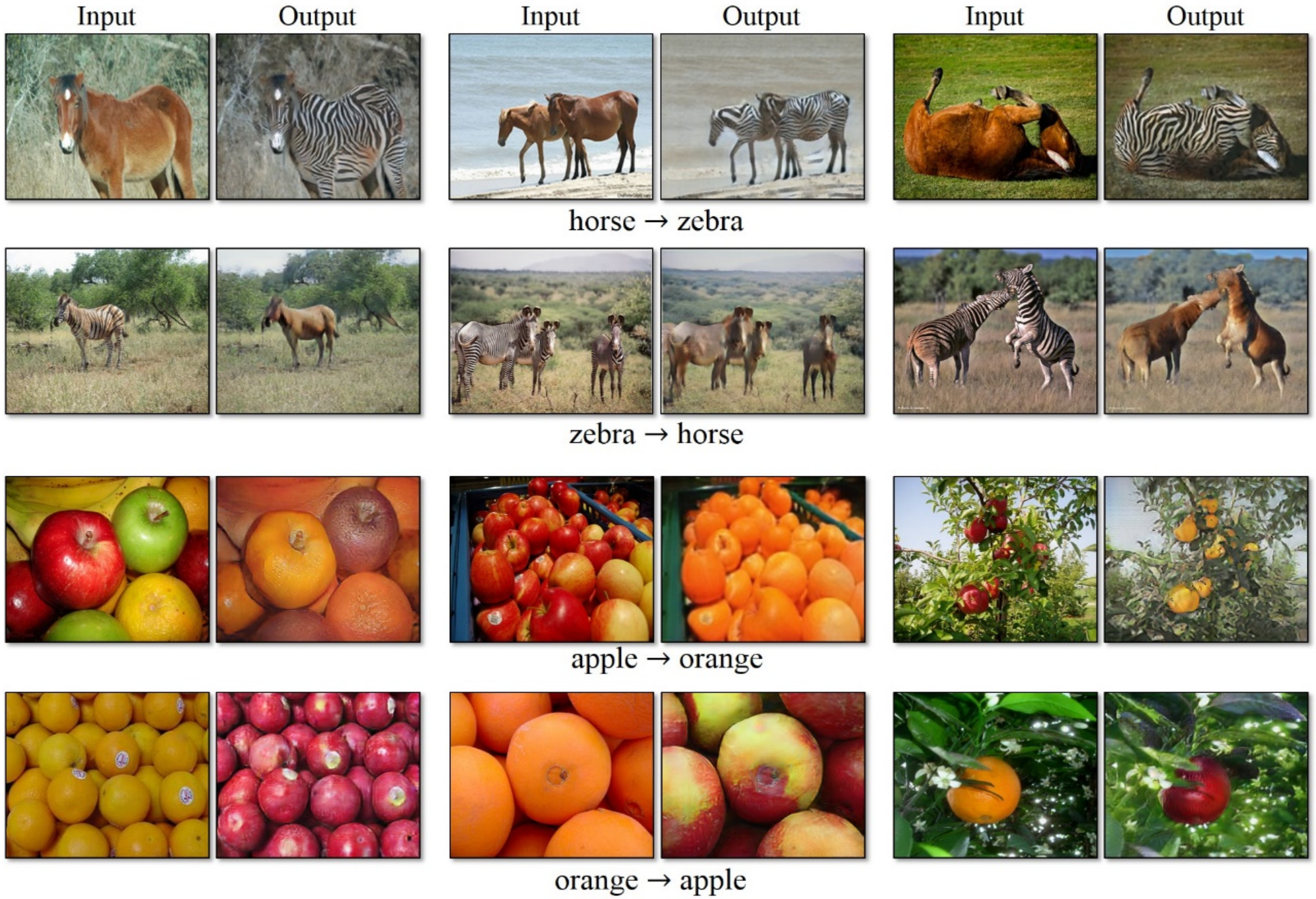
- Residual blocks
- Instance normalization
- Generated image pool





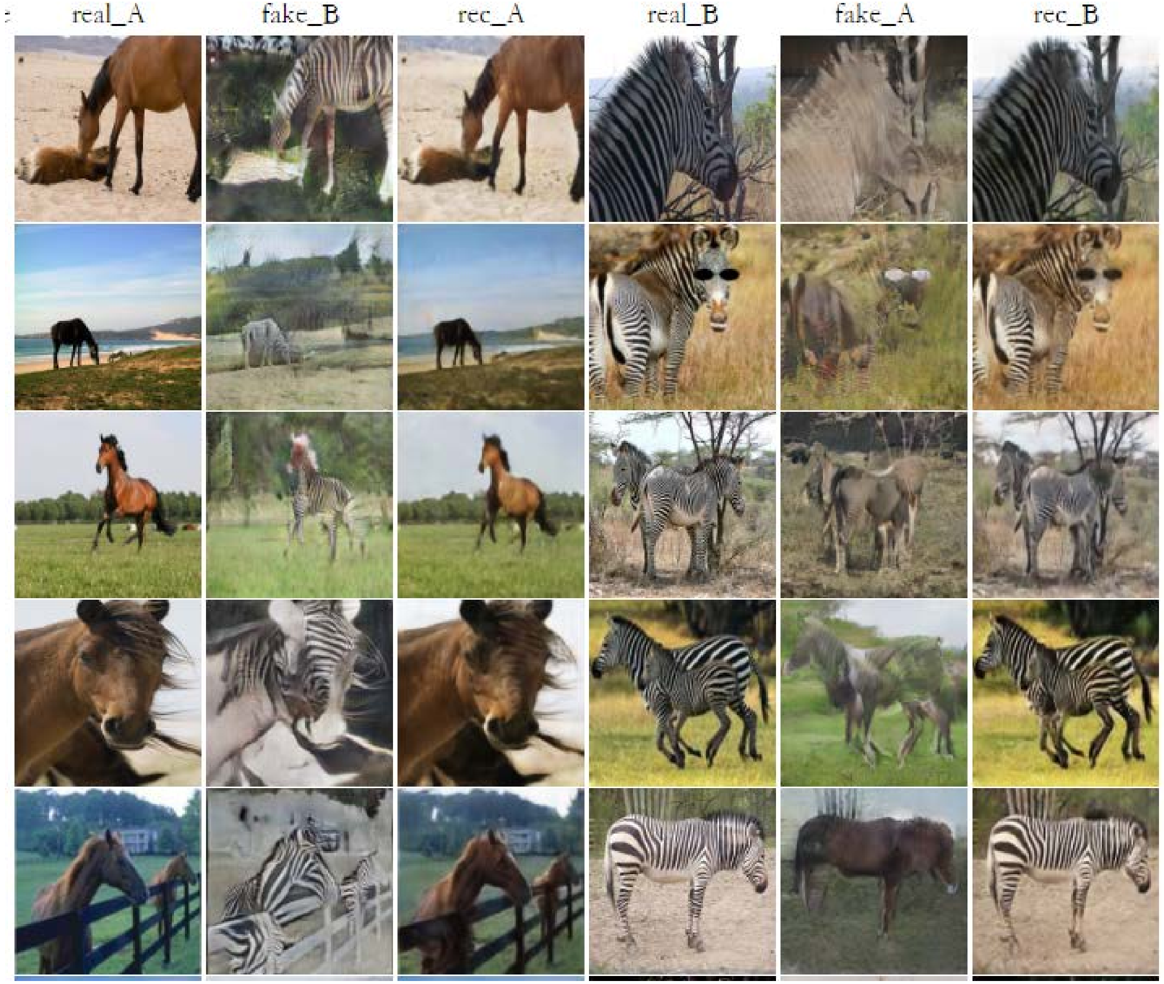
Exploration: CycleGAN

Reported results



Exploration: CycleGAN

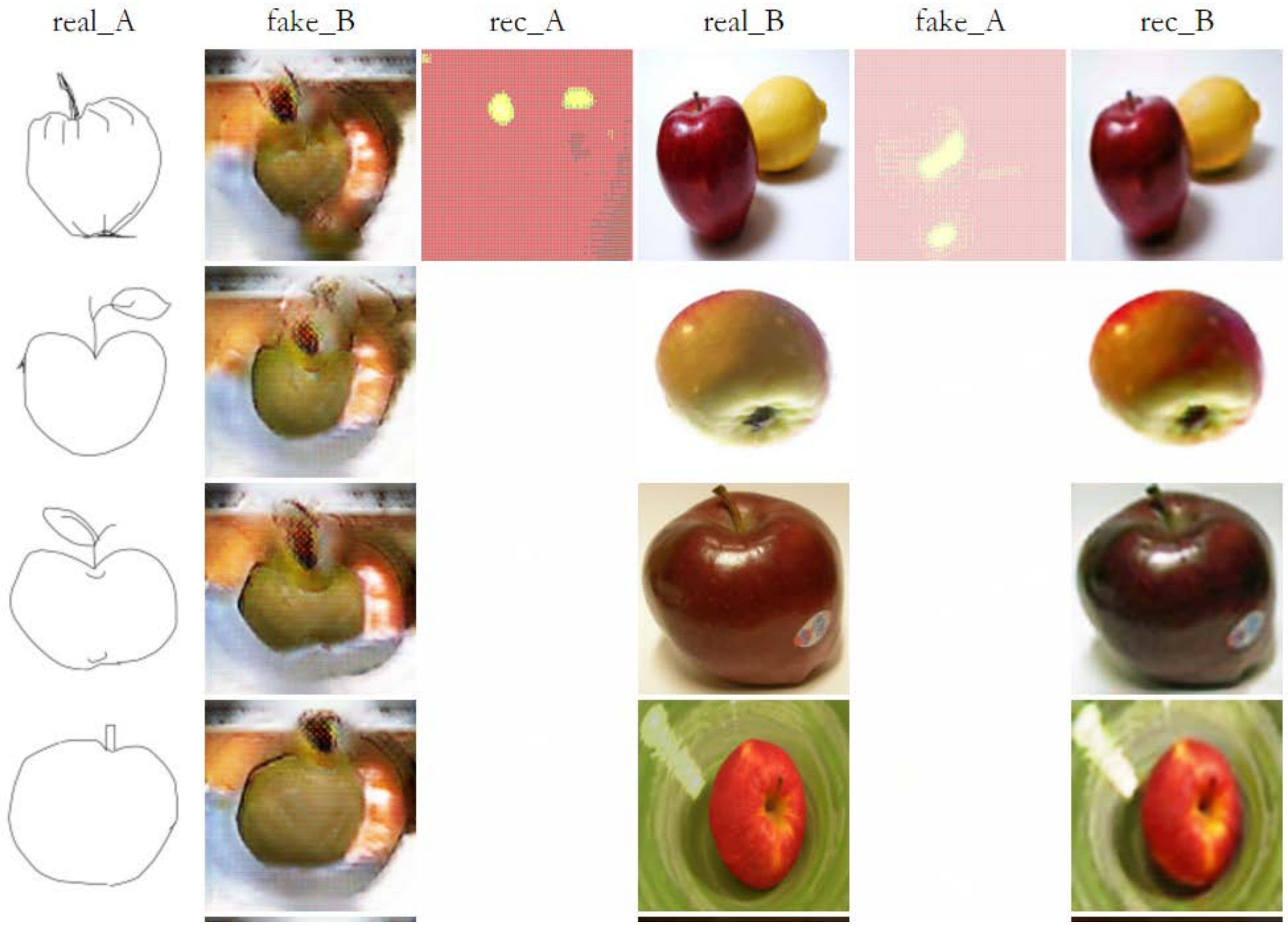
Results by training after 200 epochs with the official torch implementation. Horse to zebra dataset.



Exploration: CycleGAN

Results by training after 200 epochs with the official torch implementation. Sketch to photo (apple) dataset.

- Mode collapse (sketch to photo)
- All white (photo to sketch)



Exploration: CycleGAN

Results by training after 200 epochs with the official torch implementation. Sketch to photo (airplane) dataset.