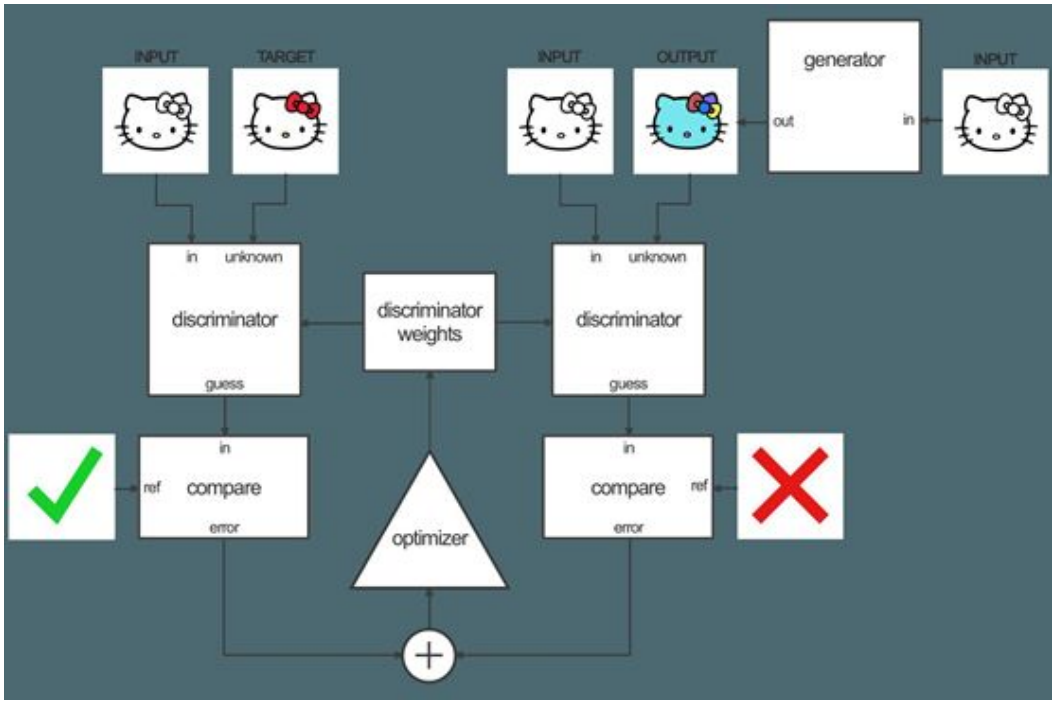
Text to Image Synthesis with Generative Adversarial Networks (GAN)

Li Cai
Jingyi Zhang
Jiahao Zhao

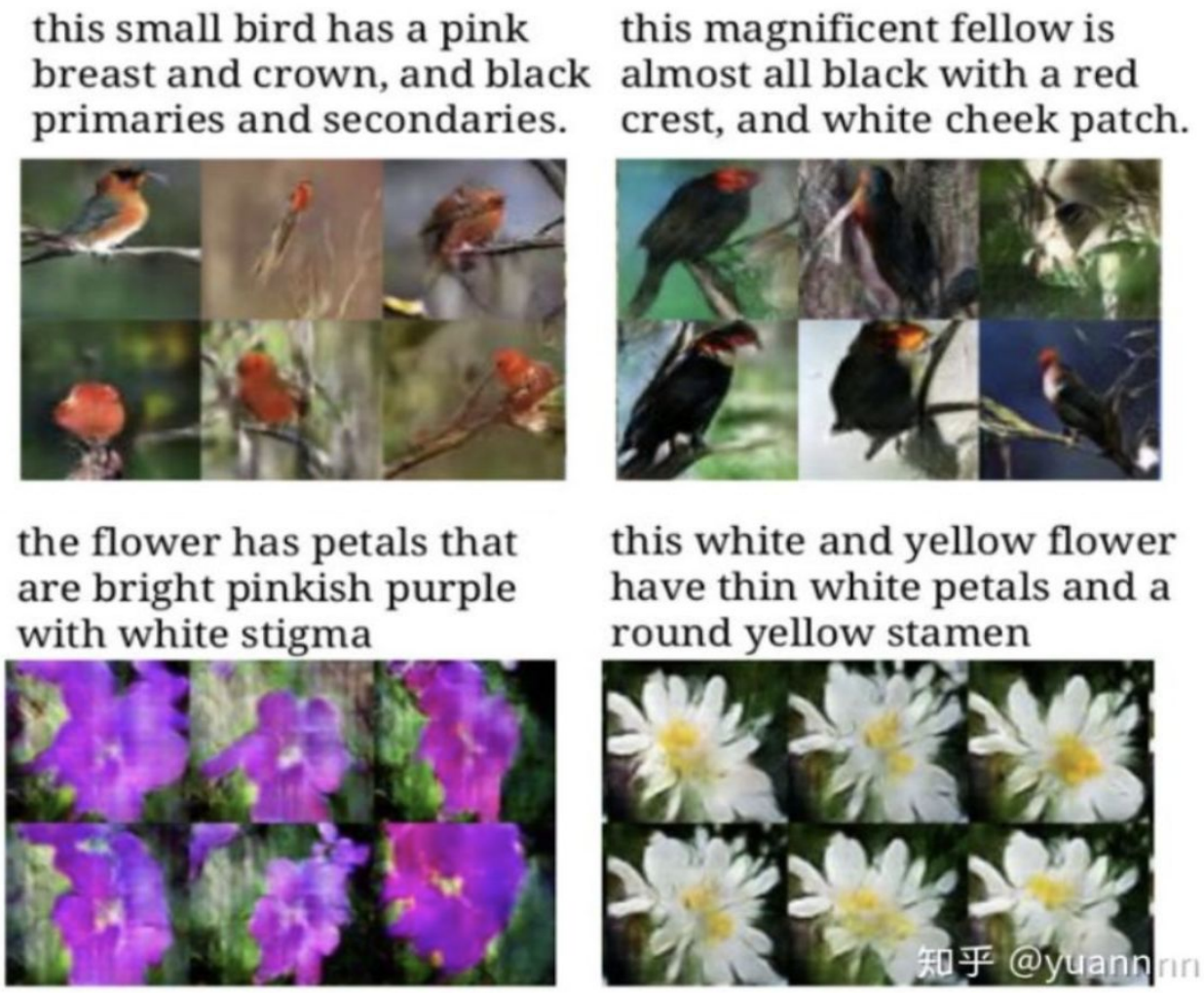
Overview

Automatic synthesis of images from text is an interesting and useful topic. With deep convolutional generative adversarial networks (GAN) and its families there has been several realization in different level.
Generally speaking, GAN is composed of two networks: the generator and the discriminator. The generator generates new stuff and passes them to the discriminator; the discriminator identifies if the one is fake.



Goals

Our work is to generate different style of images such as realistic or cartoon from one or two sentences describing how the pictures look like. Here is an example:



We plan to try some different GANs:

1) DC-GAN
2) GAWWN (Generative Adversarial What-Where Network)
3) StackGAN

Use Cases

- Accelerating the performance of image search engines.
- Artistic materials for superflat artists.
- Evaluation metrics for Object Detection Models.

Data

Image Data:

http://www.robots.ox.ac.uk/~vgg/data/flowers/102/102flowers.tgz
https://www.kaggle.com/c/quickdraw-doodle-recognition

Descriptive Text Data:

https://drive.google.com/file/d/0B0ywwgffWnLLcms2WWJQRFNSWXM/

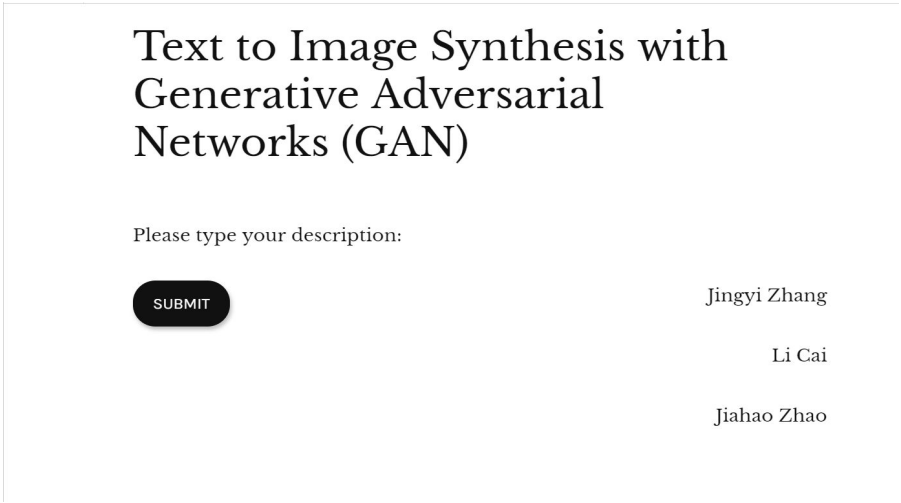
Milestones

Timeframe	Delivery
Day 1-2	Data Preprocessing and Exploratory Data Analysis
Day 3-6	Model Building, Training, Selection
Day 7-8	Deployment of models on cloud and build web application
Day 9-10	System integration and documentation

Deployment Details

1) Language: Python
2) Model Framework: Keras
3) Web Framework: Flask
4) Database: MongoDB
5) Cloud: AWS

USER INTERFACE DESIGN



Reference and Sources:

https://github.com/dashayushman/TAC-GAN
https://arxiv.org/pdf/1703.06412.pdf
http://www.scottreed.info/files/nips2016.pdf
https://arxiv.org/pdf/1605.05396.pdf