Text to Image Algorithm Understanding and Results Recurrence

Hongzhi Liu

September 28, 2018

CONTENTS

Research Purposes

1. Text-to-Image Problem

Generating images according to natural language descriptions with text format is a fundamental problem in many applications, such as art generation, computer-aided design and data sets expansion.

Most recently proposed text-to-image synthesis methods are based on Generative Adversarial Networks (GANs). AttnGAN, as the state-of-the-art algorithm, shows the best inception score on CUB and COCO test sets which are 4.36 and 25.89 respectively.

Tao Xu, et al. AttnGAN: Fine-grained text to image generation with attentional generative adversarial networks. In *CVPR*, 2018.

Research Purposes

2. My Research Goal

I should learn about AttnGAN algorithm and its architecture as shown in Fig. 1.

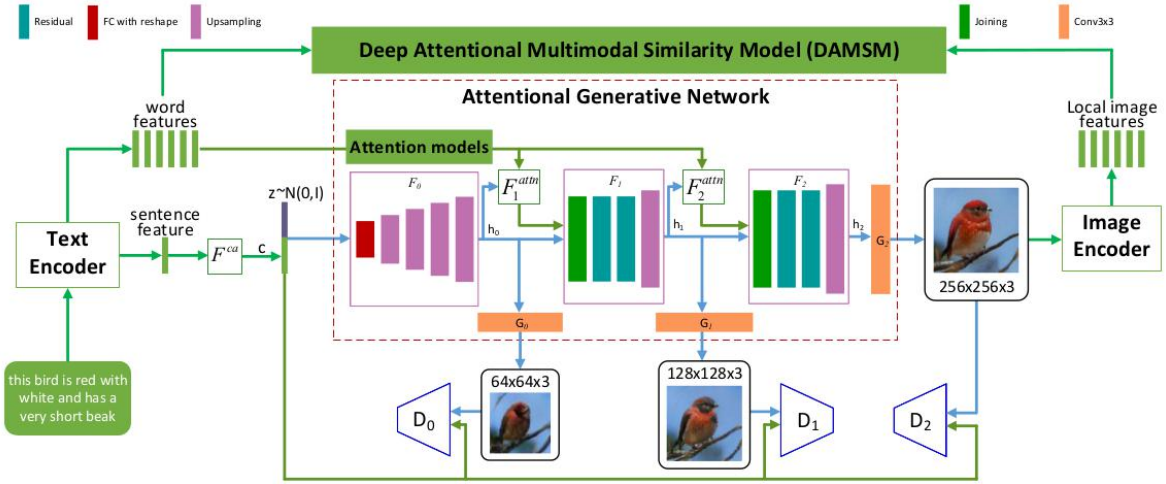

Figure1: The architecture of the proposed AttnGAN. Each attention model automatically retrieves the conditions (i.e., the most relevant word vectors) for generating different sub-regions of the image; the DAMSM provides the fine-grained image-text matching loss for the generative network.

3. Results Recurrence

I should make inception scores by GAN models reappear as shown in the table below in two weeks.

Dataset	GAWWN	StackGAN	StackGAN-v2	AttnGAN
CUB	3.62 ± .07	3.70 ± .04	3.84 ± .06	4.36 ± .03
COCO	/	8.45 ± .03	/	25.89 ± .47

S. Reed, et al. Generative adversarial text-to-image synthesis. In *ICML*, 2016.

Han Zhang, et al. StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks. In *ICCV*, 2017.

Han Zhang, et al. StackGAN++: Realistic image synthesis with stacked generative adversarial networks. *IEEE TPAMI*, 2018

Code Implementation

1. Methods

Firstly, I build runtime environment of algorithms in the server.

Next, after solving the problem of operation, I need to run the python or lua codes in order to make the algorithm generate the pictures we need including CUB and COCO pictures.

At last, I should modify the codes of computing inception score which auther provides, because the program can not directly calculate the value of scores.

Code Implementation

2. Generate CUB Pictures

The bird has gray crown, belly and white abdomen, with black tarsus and feet.

Figure2: Example CUB pictures with one sentence of the proposed AttnGAN.

Code Implementation

3. Generate COCO Pictures

A large red and white boat floating on top of a lake.

Figure3: Example COCO pictures with one sentence of the proposed AttnGAN.

Experimental Results

1. Inception Scores

Dataset	GAWWN	StackGAN	StackGAN-v2	AttnGAN
CUB	3.62 ± .07	3.70 ± .04	3.84 ± .06	4.36 ± .03
Ours	2.05 ± .23	3.91± .03	4.10± .04	4.24 ± .18
COCO	/	8.45 ± .03	/	25.89 ± .47
Ours	/	9.74±.50	/	23.96± .35

Table1: Inception scores by state-of-the art GAN models and my results recurrence.

Further Research

1. SSD evaluation

2. ResNeXt algorithm and OpenGL GLU

3. Multiple target generation

Q & A