# Text To Image Synthesis Using Thought Vectors

[![Join the chat at https://gitter.im/text-to-image/Lobby](https://badges.gitter.im/text-to-image/Lobby.svg)](https://gitter.im/text-to-image/Lobby?utm\_source=badge&utm\_medium=badge&utm\_campaign=pr-badge&utm\_content=badge)

This is an experimental tensorflow implementation of synthesizing images from captions using [Skip Thought Vectors][1]. The images are synthesized using the GAN-CLS Algorithm from the paper [Generative Adversarial Text-to-Image Synthesis][2]. This implementation is built on top of the excellent [DCGAN in Tensorflow][3]. The following is the model architecture. The blue bars represent the Skip Thought Vectors for the captions.

![Model architecture](http://i.imgur.com/dNl2HkZ.jpg)

Image Source : [Generative Adversarial Text-to-Image Synthesis][2] Paper

## Requirements

- Python 2.7.6

- [Tensorflow][4]

- [h5py][5]

- [Theano][6] : for skip thought vectors

- [scikit-learn][7] : for skip thought vectors

- [NLTK][8] : for skip thought vectors

## Datasets

- All the steps below for downloading the datasets and models can be performed automatically by running `python download\_datasets.py`. Several gigabytes of files will be downloaded and extracted.

- The model is currently trained on the [flowers dataset][9]. Download the images from [this link][9] and save them in ```Data/flowers/jpg```. Also download the captions from [this link][10]. Extract the archive, copy the ```text\_c10``` folder and paste it in ```Data/flowers```.

- Download the pretrained models and vocabulary for skip thought vectors as per the instructions given [here][13]. Save the downloaded files in ```Data/skipthoughts```.

- Make empty directories in Data, ```Data/samples```, ```Data/val\_samples``` and ```Data/Models```. They will be used for sampling the generated images and saving the trained models.

## Usage

- <b>Data Processing</b> : Extract the skip thought vectors for the flowers data set using :

```

python data\_loader.py --data\_set="flowers"

```

- <b>Training</b>

\* Basic usage `python train.py --data\_set="flowers"`

\* Options

- `z\_dim`: Noise Dimension. Default is 100.

- `t\_dim`: Text feature dimension. Default is 256.

- `batch\_size`: Batch Size. Default is 64.

- `image\_size`: Image dimension. Default is 64.

- `gf\_dim`: Number of conv in the first layer generator. Default is 64.

- `df\_dim`: Number of conv in the first layer discriminator. Default is 64.

- `gfc\_dim`: Dimension of gen untis for for fully connected layer. Default is 1024.

- `caption\_vector\_length`: Length of the caption vector. Default is 1024.

- `data\_dir`: Data Directory. Default is `Data/`.

- `learning\_rate`: Learning Rate. Default is 0.0002.

- `beta1`: Momentum for adam update. Default is 0.5.

- `epochs`: Max number of epochs. Default is 600.

- `resume\_model`: Resume training from a pretrained model path.

- `data\_set`: Data Set to train on. Default is flowers.

- <b>Generating Images from Captions</b>

\* Write the captions in text file, and save it as ```Data/sample\_captions.txt```. Generate the skip thought vectors for these captions using:

```

python generate\_thought\_vectors.py --caption\_file="Data/sample\_captions.txt"

```

\* Generate the Images for the thought vectors using:

```

python generate\_images.py --model\_path=<path to the trained model> --n\_images=8

```

```n\_images``` specifies the number of images to be generated per caption. The generated images will be saved in ```Data/val\_samples/```. ```python generate\_images.py --help``` for more options.

## Sample Images Generated

Following are the images generated by the generative model from the captions.

| Caption | Generated Images |

| ------------- | -----:|

| the flower shown has yellow anther red pistil and bright red petals | ![](http://i.imgur.com/SknZ3Sg.jpg) |

| this flower has petals that are yellow, white and purple and has dark lines | ![](http://i.imgur.com/8zsv9Nc.jpg) |

| the petals on this flower are white with a yellow center | ![](http://i.imgur.com/vvzv1cE.jpg) |

| this flower has a lot of small round pink petals. | ![](http://i.imgur.com/w0zK1DC.jpg) |

| this flower is orange in color, and has petals that are ruffled and rounded. | ![](http://i.imgur.com/VfBbRP1.jpg) |

| the flower has yellow petals and the center of it is brown | ![](http://i.imgur.com/IAuOGZY.jpg) |

## Implementation Details

- Only the uni-skip vectors from the skip thought vectors are used. I have not tried training the model with combine-skip vectors.

- The model was trained for around 200 epochs on a GPU. This took roughly 2-3 days.

- The images generated are 64 x 64 in dimension.

- While processing the batches before training, the images are flipped horizontally with a probability of 0.5.

- The train-val split is 0.75.

## Pre-trained Models

- Download the pretrained model from [here][14] and save it in ```Data/Models```. Use this path for generating the images.

## TODO

- Train the model on the MS-COCO data set, and generate more generic images.

- Try different embedding options for captions(other than skip thought vectors). Also try to train the caption embedding RNN along with the GAN-CLS model.

## References

- [Generative Adversarial Text-to-Image Synthesis][2] Paper

- [Generative Adversarial Text-to-Image Synthesis][11] Code

- [Skip Thought Vectors][1] Paper

- [Skip Thought Vectors][12] Code

- [DCGAN in Tensorflow][3]

- [DCGAN in Tensorlayer][15]

## Alternate Implementations

- [Text to Image in Torch by Scot Reed][11]

- [Text to Image in Tensorlayer by Dong Hao][16]

## License

MIT

[1]:http://arxiv.org/abs/1506.06726

[2]:http://arxiv.org/abs/1605.05396

[3]:https://github.com/carpedm20/DCGAN-tensorflow

[4]:https://github.com/tensorflow/tensorflow

[5]:http://www.h5py.org/

[6]:https://github.com/Theano/Theano

[7]:http://scikit-learn.org/stable/index.html

[8]:http://www.nltk.org/

[9]:http://www.robots.ox.ac.uk/~vgg/data/flowers/102/

[10]:https://drive.google.com/file/d/0B0ywwgffWnLLcms2WWJQRFNSWXM/view

[11]:https://github.com/reedscot/icml2016

[12]:https://github.com/ryankiros/skip-thoughts

[13]:https://github.com/ryankiros/skip-thoughts#getting-started

[14]:https://bitbucket.org/paarth\_neekhara/texttomimagemodel/raw/74a4bbaeee26fe31e148a54c4f495694680e2c31/latest\_model\_flowers\_temp.ckpt

[15]:https://github.com/zsdonghao/dcgan

[16]:https://github.com/zsdonghao/text-to-image